

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 03:21:30 ; Search time 669 Seconds

(without alignments)
10172.815 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacacgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_29Jan04:*

- 1: Geneseq19806:*
- 2: Geneseq19906:*
- 3: Geneseq20006:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	100.0	1602	4	AAF74206
2	1138.6	71.1	2235	5	AAD16797
3	676.2	42.2	1962	4	AAH44262
4	621.2	38.8	1120	5	AAH46798
5	350.8	21.9	636	8	ACL18332
6	324.6	20.3	605	8	ACL25710
7	304.2	19.0	566	8	ACL19398
8	303.6	19.0	487	9	ADH82005
9	302.8	18.9	611	8	ACL25711
10	299.8	18.7	777	8	ACL19405
11	280.6	17.5	575	8	ACL18347
12	276.6	17.3	667	8	ACL18333
13	271.8	17.0	622	8	ACL18325
14	268.6	16.8	589	8	ACL25708
15	255.8	16.0	619	4	AAH44246
16	248.4	15.5	650	8	ACL18342
17	248.2	15.5	736	8	ACL18395
18	240.8	15.0	671	8	ACL19401
19	240.2	15.0	738	8	ACL18354
20	240.0	15.0	517	8	ACL18349
21	239.4	14.9	436	8	ACL18323
22	237.2	14.8	712	8	ACL18373
23	235.2	14.7	663	8	ACL19412

C 24	233.4	14.6	648	8	ACL19403
C 25	233.4	14.6	676	8	ACL18366
C 26	231.2	14.4	727	8	ACL18362
C 27	229.8	14.3	477	8	ACL25703
C 28	227	14.2	621	8	ACL18331
C 29	227	14.2	659	8	ACL18389
C 30	224.6	14.0	696	8	ACL18382
C 31	224.2	14.0	649	8	ACL19408
C 32	224.2	14.0	652	8	ACL19407
C 33	223.6	14.0	624	8	ACL19409
C 34	222.2	13.9	623	8	ACL18326
C 35	222.2	13.9	681	8	ACL18374
C 36	221.6	13.8	697	8	ACL18372
C 37	220.2	13.7	692	8	ACL18391
C 38	213	13.3	587	8	ACL18344
C 39	212.8	13.3	527	8	ACL18334
C 40	211.6	13.2	695	8	ACL18380
C 41	211.2	13.2	684	8	ACL18377
C 42	209.6	13.1	631	8	ACL18356
C 43	209.6	13.1	673	8	ACL18378
C 44	204.8	12.8	617	8	ACL19410
C 45	203.2	12.7	683	8	ACL18383

ALIGNMENTS

RESULT 1

AAF74206

ID AAF74206 standard; DNA; 1602 BP.

XX AC AAF74206;

XX DT 06-AUG-2003 (revised)

XX DT 02-MAY-2001 (first entry)

XX DE DNA encoding environmental stress tolerant protein SEQ ID 39.

XX KM Environmental stress resistance; salt; heat; desert; transgenic plant;

XX KM ds.

XX OS Suaeda japonica.

XX PN WC200106006-AI.

XX PD 25-JAN-2001.

XX PF 19-JUL-2000; 2000WO-JP004862.

XX PR 19-JUL-1999; 99JP-00235910.

XX PR 24-MAR-2000; 2000JP-00085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Yamada A, Ozeki Y, Saito T;

XX DR WPI; 2001-147355/15.

XX P-PSDB; AAB80627.

XX Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

XX Claim 65; Page 123-125; 167pp; Japanese.

XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell. The host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 03:21:30 ; Search time 669 Seconds
(without alignments)
10172.815 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacaccgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_25Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1602	100.0	1602	4	Aaf74206 DNA encod
2	1138.6	71.1	2235	5	Aad16797 Spinach P
3	676.2	42.2	1962	4	Aah44262 Physcomit
4	621.2	38.8	1120	5	Aad16798 Spinach P
5	350.8	21.9	636	8	ACL18332 DNA clone
6	324.6	20.3	605	8	ACL25710 DNA clone
7	304.2	19.0	566	8	ACL19398 DNA clone
8	303.6	19.0	487	9	Ade82005 Arabidops
9	302.8	18.9	611	8	ACL25711 DNA clone
10	299.8	18.7	777	8	ACL19405 DNA clone
11	280.6	17.5	575	8	ACL18347 DNA clone
12	276.6	17.3	667	8	ACL18333 DNA clone
13	271.8	17.0	622	8	ACL18325 DNA clone
14	268.6	16.8	589	8	ACL25708 DNA clone
15	255.8	16.0	619	4	Aah44246 Physcomit
16	248.4	15.5	650	8	ACL18342 DNA clone
17	248.2	15.5	736	8	ACL18395 DNA clone
18	240.8	15.0	671	8	ACL19401 DNA clone
19	240.2	15.0	738	8	ACL18354 DNA clone
20	240.0	15.0	517	8	ACL18349 DNA clone
21	239	14.9	436	8	ACL18323 DNA clone
22	237.2	14.8	712	8	ACL18373 DNA clone
23	235.2	14.7	663	8	ACL19412 DNA clone

C 24	233.4	14.6	648	8	ACL19403	ACL19403 DNA clone
C 25	233.4	14.6	676	8	ACL18366	ACL18366 DNA clone
C 26	231.2	14.4	727	8	ACL18362	ACL18362 DNA clone
C 27	229.8	14.3	477	8	ACL25703	ACL25703 DNA clone
C 28	227	14.2	621	8	ACL18331	ACL18331 DNA clone
C 29	227	14.2	659	8	ACL18389	ACL18389 DNA clone
C 30	224.6	14.0	696	8	ACL18382	ACL18382 DNA clone
C 31	224.2	14.0	649	8	ACL19408	ACL19408 DNA clone
C 32	224.2	14.0	652	8	ACL19407	ACL19407 DNA clone
C 33	223.6	14.0	624	8	ACL19409	ACL19409 DNA clone
C 34	222.2	13.9	623	8	ACL18326	ACL18326 DNA clone
C 35	222.2	13.9	681	8	ACL18374	ACL18374 DNA clone
C 36	221.6	13.8	697	8	ACL18372	ACL18372 DNA clone
C 37	220.2	13.7	692	8	ACL18391	ACL18391 DNA clone
C 38	213	13.3	587	8	ACL18344	ACL18344 DNA clone
C 39	212.8	13.3	527	8	ACL18334	ACL18334 DNA clone
C 40	211.6	13.2	695	8	ACL18380	ACL18380 DNA clone
C 41	211.2	13.2	684	8	ACL18377	ACL18377 DNA clone
C 42	209.6	13.1	631	8	ACL18356	ACL18356 DNA clone
C 43	209.6	13.1	673	8	ACL18378	ACL18378 DNA clone
C 44	204.8	12.8	617	8	ACL19410	ACL19410 DNA clone
C 45	203.2	12.7	583	8	ACL18383	ACL18383 DNA clone

ALIGNMENTS

RESULT 1

AAF74206
ID AAF74206 standard; DNA; 1602 BP.
XX
AC AAF74206;
XX
DT 06-AUG-2003 (revised)
DT 02-MAY-2001 (first entry)
XX
DE DNA encoding environmental stress tolerant protein SEQ ID 39.
XX
KW Environmental stress resistance; salt; heat; desert; transgenic plant;
KW ds.
XX
OS Suaeda japonica.
XX
PN WO200106006-A1.
XX
PD 25-JAN-2001.
XX
PF 19-JUL-2000; 2000WO-JP004862.
XX
PR 19-JUL-1999; 99JP-00235910.
PR 24-MAR-2000; 2000JP-00085377.
XX
(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Yamada A, Ozeki Y, Saito T;
XX
XX WPI; 2001-147355/15.
DR P-PSDB; AAB80627.
XX
XX Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.
XX
XX Claim 65; Page 123-125; 167pp; Japanese.
XX
XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed cells does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance

CC factors. The DNA encoding proteins conferring environmental stress
CC resistance, can be used in the production of plants resistant to
CC environmental stress, which can be cultivated in unfavourable
CC environments such as deserts, salt damaged ground, cold regions and the
CC oceans. They can be used for increasing the area of land covered by green
CC plants, and desert greening and afforestation, in order to counter the
CC effects of the increase in atmospheric carbon dioxide concentration. PCR
CC primers AAF74219 and AAP74220 are used in an example illustrating the
CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)
XX

SQ Sequence 1602 BP; 454 A; 279 C; 383 G; 486 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1602;	DB 4;	Length 1602;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1602;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CACCGCTGATTAACCAATGAAGCTATGATGCTCGATTCCTCAAGCTTCTCATCTTCAC	60	
Db	1	CACCGCTGATTAACCAATGAAGCTATGATGCTCGATTCCTCAAGCTTCTCATCTTCAC	60	
Qy	61	AAAGAGAAAGCTGCTGAGATTTCTTCAATGCTTCGCGCTCTTGAAGAAATGCTCTTG	120	
Db	61	AAAGAGAAAGCTGCTGAGATTTCTTCAATGCTTCGCGCTCTTGAAGAAATGCTCTTG	120	
Qy	121	GAACTGGGCGCTGATGCTGCTGCTTACTGCTGAATGGCTGAGAAAGCTGCGCAGGT	180	
Db	121	GAACTGGGCGCTGATGCTGCTGCTTACTGCTGAATGGCTGAGAAAGCTGCGCAGGT	180	
Qy	181	ATTGCTCTGGATTTCAATGAGTGTCTATCAGAGGATCAAGTATCAATGAGGCACTAC	240	
Db	181	ATTGCTCTGGATTTCAATGAGAGTGTCTATCAGAGGATCAAGTATCAATGAGGCACTAC	240	
Qy	241	AAAGATGTCAGTTTATGCTGCTGATGCTTCTCCACTCTCAGTTTCCACACAT	300	
Db	241	AAAGATGTCAGTTTATGCTGCTGATGCTTCTCCACTCTCAGTTTCCACACAT	300	
Qy	301	TCAATGGATGATATCTCCAAATGGTGTACTCATGTATCTTCTCATGAGAGGTGAA	360	
Db	301	TCAATGGATGATATCTCCAAATGGTGTACTCATGTATCTTCTCATGAGAGGTGAA	360	
Qy	361	ATTGCTGTAAGAAATGTTGAAATGGTGTGAAGCAGCGGGTTACATTTCTTCAGAA	420	
Db	361	ATTGCTGTAAGAAATGTTGAAATGGTGTGAAGCAGCGGGTTACATTTCTTCAGAA	420	
Qy	421	TCTGTGTTCCATCAATCTGGGGATCAAAAGCGCAAAAGCAATCCCACTACCGTGAA	480	
Db	421	TCTGTGTTCCATCAATCTGGGGATCAAAAGCGCAAAAGCAATCCCACTACCGTGAA	480	
Qy	481	CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCAATTCAGATGATGATGGAATCTCT	540	
Db	481	CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCAATTCAGATGATGATGGAATCTCT	540	
Qy	541	TATGAGCTCTCCCTACTAGCTGCAAAATGTAATGGAGCTTATGTCAGAAACAGAAAC	600	
Db	541	TATGAGCTCTCCCTACTAGCTGCAAAATGTAATGGAGCTTATGTCAGAAACAGAAAC	600	
Qy	601	CAGAACCAATGATGTTGGTGGCAAAAGTGTGATTCAGAGATGATGAGGTTCCAG	660	
Db	601	CAGAACCAATGATGTTGGTGGCAAAAGTGTGATTCAGAGATGATGAGGTTCCAG	660	
Qy	661	CGATTTCTGGATCTAGCCAGTCAAGTGTATAGCATTCGCGATATGAGCGTATTT	720	
Db	661	CGATTTCTGGATCTAGCCAGTCAAGTGTATAGCATTCGCGATATGAGCGTATTT	720	
Qy	721	GGCCCTGTTATGTTAGCACTGGAGGATATGAACACCAAGAGTTTGTCTCATGCTG	780	
Db	721	GGCCCTGTTATGTTAGCACTGGAGGATATGAACACCAAGAGTTTGTCTCATGCTG	780	
Qy	781	GACTGAGCTGGCGAGAGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	840	
Db	781	GACTGAGCTGGCGAGAGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT	840	
Qy	841	TACATGCGGAGACCTTTGATGCTGAGGTTGTTGATGCTCTCGTTATATGATTT	900	

Db	841	TACATGCGGAGACCTTTGATGCTGAGGTTGTTGATGCTCTCCGTTATATGATTT	900	
Qy	901	TCTTTTGGCTTGGAGCTTCTATTTGGGCTTAAATGCTGCTTGAAGTTGAGGTAGCAGAT	960	
Db	901	TCTTTTGGCTTGGAGCTTCTATTTGGGCTTAAATGCTGCTTGAAGTTGAGGTAGCAGAT	960	
Qy	961	TGCACCAAGATTAATCACTACCTGATTAATCTTTTGTGATCTATAGGCGTGACACCAT	1020	
Db	961	TGCACCAAGATTAATCACTACCTGATTAATCTTTTGTGATCTATAGGCGTGACACCAT	1020	
Qy	1021	CTGCATATTCAGCAAGACCTGCTGCTTGTAGATCTCTTACAAATGTTGAAGCCAGCA	1080	
Db	1021	CTGCATATTCAGCAAGACCTGCTGCTTGTAGATCTCTTACAAATGTTGAAGCCAGCA	1080	
Qy	1081	GCTAAAGTTCTAATCAGTGAATTAATGCAAGAAAGCTGGTCCACCTCAGTGAATTCGCC	1140	
Db	1081	GCTAAAGTTCTAATCAGTGAATTAATGCAAGAAAGCTGGTCCACCTCAGTGAATTCGCC	1140	
Qy	1141	GCTTACATTAAGCAGAGGGATATGATCTCCATGATGTAAGGATATGGCGAGATGCTT	1200	
Db	1141	GCTTACATTAAGCAGAGGGATATGATCTCCATGATGTAAGGATATGGCGAGATGCTT	1200	
Qy	1201	AAAGATGCTGCAATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTTCAATGAGTT	1260	
Db	1201	AAAGATGCTGCAATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTTCAATGAGTT	1260	
Qy	1261	CTACCGAAGGAACTAGAGACTGTTGAGAGCAAAAGGATGTTTCAATGATGATTTCTCT	1320	
Db	1261	CTACCGAAGGAACTAGAGACTGTTGAGAGCAAAAGGATGTTTCAATGATGATTTCTCT	1320	
Qy	1321	GAGGAGGATTAATGATGATTTGTTGAGGTTGAGATGATGATGTTGCGAGGAGTCCAG	1380	
Db	1321	GAGGAGGATTAATGATGATTTGTTGAGGTTGAGATGATGATGTTGCGAGGAGTCCAG	1380	
Qy	1381	GGTGAGCAACGATGGGGTCTGTTGTTGCCAAGAAAGGATGAAAGATCAGTTGCGCACCT	1440	
Db	1381	GGTGAGCAACGATGGGGTCTGTTGTTGCCAAGAAAGGATGAAAGATCAGTTGCGCACCT	1440	
Qy	1441	GGCAGTGTGATTTCTTCTAGTATTAATCTTCAATGTTTCTCATGATGATCTCTACATGT	1500	
Db	1441	GGCAGTGTGATTTCTTCTAGTATTAATCTTCAATGTTTCTCATGATGATCTCTACATGT	1500	
Qy	1501	AAATTTGCCAATTAAGTTCATTTGCGAGATGTTGAAGATGATTAATCATATTTTATCTTTT	1560	
Db	1501	AAATTTGCCAATTAAGTTCATTTGCGAGATGTTGAAGATGATTAATCATATTTTATCTTTT	1560	
Qy	1561	AATTAATCATGATTTATGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	1602	
Db	1561	AATTAATCATGATTTATGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	1602	

RESULT 2

AAD16797
ID AAD16797 standard; DNA; 2235 BP.

AC AAD16797;

DT 29-NOV-2001 (first entry)

DE Spinach PEAMT DNA.

Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;
PEAMT; cellular intermediate; phospho-dimethylethanolamine, choline;
phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
glycine betaine; choline-O-sulphate; lipid content alteration;
osmotic stress tolerance; nutritional value; transgenic plant;
cryoprotectant; ds.

OS Spinacia oleracea.

Key Location/Qualifiers

FT CDS 254..1738

Db 1552 TTGCTGATACATTACGACAGAGGGTATGATCTCCATAGCTTCAGAGTACGGAG 1611
 QY 1195 ATGCTTAAGATGCTGGATTGTTGATCTTCCGAGGATAGAACTGACGATTCNTT 1254
 Db 1612 ATCTGAGATGCCGTTTGTGGAGTGTGCGACAGGACCCGACGATCAGTTCAT 1671
 QY 1255 CGAGTCTACGAGGAACTAGAGACTGTTGAGAGGAAAGAGATGTTTCATAGTAT 1314
 Db 1672 GAGATGTTACAGAGGAGTACCCACCTGAAGCAGGTGTCACCAAGAT 1731
 QY 1315 TTCTCTGAGGAGATTAACATGATCTGTTGAGGTTGGAATGATAAGTTGCGGAGACT 1374
 Db 1732 TTCTCCGAGGAGATTAACATGATCTGTTGAGGAGTGAAGAGTGAAGCTGAGGCGTGT 1791
 QY 1375 GCCAGGCTGACGACGATGGGCTGTTGCTGTCGACAGAGAGTGAAGA 1425
 Db 1792 TCGATGACGACAGAGTGGGACTCTTCATAGCTTACAGGCAATATGA 1842

RESULT 4

AADI6798

ID AADI6798 standard; DNA; 1120 BP.

XX AC AADI6798;

XX DT 29-NOV-2001 (first entry)

XX DE Spinach PEAMT truncated DNA.

XX KW Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;

XX KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;

XX KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;

XX KW glycine betaine; choline-O-sulphate; lipid content alteration;

XX KW osmotic stress tolerance; nutritional value; transgenic plant;

XX KW cryoprotectant; ds.

XX OS Spinacia oleracea.

XX PH Key

XX FT Location/Qualifiers

XX FT CDS 254..1120

XX FT /tag= a

XX FT /product= "Spinach PEAMT truncated protein"

XX FT /note= "CDS does not include stop codon"

XX FT /partial

XX PW WO200168870-A2.

XX PD 20-SEP-2001.

XX PF 15-MAR-2001; 2001WO-US008352.

XX PR 15-MAR-2000; 2000US-00525885.

XX PA (UYEL) UNIV FLORIDA.

XX PA (UYCA-) UNIV CARNEGIE MELLON.

XX PI Hanson AD, Nuccio ML, Henry SA;

XX DR WPI; 2001-565796/63.

XX DR P-PSDB; AAE09761.

XX XX

XX PT New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase

XX PT polypeptides, useful for modulating the levels of cellular intermediates

XX PT such as phosphodimethylethanolamine and for altering the lipid content in

XX PT plants cells.

XX PS Claim 18; Page 109; 158pp; English.

XX CC The present sequence is: spinach S-adenosyl-L-methionine:phospho-

XX CC ethanolamine N-methyltransferase (PEAMT) truncated DNA. The PEAMT

XX CC sequences are useful for modulating the levels of cellular intermediates

XX CC such as phosphodimethylethanolamine, phosphomono-methylethanolamine,

CC choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or
 CC glycine betaine. They are useful for altering the lipid content in plant
 CC cells. The polynucleotides are also useful for improving the osmotic
 CC stress tolerance of a plant and increasing the cryoprotectant properties
 CC of a plant. The present invention also relates to methods and
 CC compositions comprising PEAMT used for generating transgenic plants with
 CC increased nutritional value

SQ Sequence 1120 BP; 304 A; 218 C; 256 G; 342 T; 0 U; 0 Other;

Query Match 38.8%; Score 621.2; DB 5; Length 1120;

Best Local Similarity 86.4%; Pred. No. 3.1e-147;

Matches 686; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 CACACCTGCTGATTAACCACTGAAGCTATGATGCTCGATTTCTCAAGCTTCGTGATCTTGAC 60
 Db 317 CACTCTGTTGATTTGACTGTGTGAGGCTATGATGCTGATTCACAAGCTTCAGATCTTGAC 376
 QY 61 AAGAAGAACGTCCTGAGATTTCTTCAATGCTTCCGCTCTTTGAAGGAAATGCTCTTG 120
 Db 377 AAGTGGAGCGACCTGAGGTACTTTCATGCTTCCACCTTATGAAGGAAAGCTGTGCTTA 436
 QY 121 GAACTCGGGCTGATTTGCTGCTTTTACTGCTGAATGGCTGAGAAAGCTGGCCAGGTT 180
 Db 437 GAATCGGCTGCTGATTTGCTGCTTTTACTGCTGAATGGCTGAGAAAGCTGGCCAGGTC 496
 QY 181 ATTGCTCTGATTTCAATTCAGAGTCTATCAAGAAAGATGAAGTATCAATGGGCACTAC 240
 Db 497 ATGCTCTGATTTCAATTCAGAGTGTGTATTAAGAAAGATGAAGCATTAATGGGCACTAC 556
 QY 241 AAAAAATGCTCAAGTTTATGCTGCTGATGTGACTTCTCCCACTCTCAGTTTCCCAACCAT 300
 Db 557 AAAAAATGCTCAAGTTTATGCTGCTGATGTGACTTCTCCCACTCTCAGTTTCCCAACCAT 616
 QY 301 TCATTGATGATGATTTCTCCCAATGCTTACTCATGTATCTTCTGATGAAGAGGTGAA 360
 Db 617 TCCGTGATGATGATTTCTCCCAATGCTTACTCATGTATCTTCTGATGAAGAGGTGAG 676
 QY 361 AATTGCTTGAAGAAATGTTGAATGGTTGAAGCCAGGGGGTTACATTTCTTCAGAGAA 420
 Db 677 CGTCTGGTTGAAGAGGATGTTGAATGGTTGAAGCCAGGAGGATACATTTCTTCAGAGAA 736
 QY 421 TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCAACCACTTACCCTGAA 480
 Db 737 TCTTGTTCATCAATCAGGAGATCAAAAGCGCAAAAGCAATCCCAACCACTTACCCTGAA 796
 QY 481 CTTAGGTTCTACACTAGGCTTCAAGAGTGTCTATTGCAAGATGATCTGGAACCTCT 540
 Db 797 CTTAGGTTCTACCAAGATCTTCAAGAGATGCTCATATGCAAGATGATCTGGAACCTCC 856
 QY 541 TATGAGCTCTCCCTACTTAGCTCAATGTATTGGAGCTTATGTTCAGAAACAAAGAAAC 600
 Db 857 TATGAGCTCTCCCTAAATGGCTCAATGTATTGGAGCTTATGTTCAGAAACAAAGAAAT 916
 QY 601 CAGAACAGATGATGCTGTTGCGCAAAAGTTGATTTCTTAAGATGATTAAGGGTTCAG 660
 Db 917 CAGAACAGATGATGCTGTTTATGCGCAAAAGTTGATTTCTAGAGATGATCAAGGGTTCAG 976
 QY 661 CGATTTCTGGATCTAGCCAGTCAAGTGTATAGCAATCTTCGATATGAGCGGTGATTT 720
 Db 977 CGATTTCTGGATCTAGTCAATCAAGTTTAAACAGCATACTCGGTATGAGCGGTGATTT 1036
 QY 721 GGCCTCTGTTATGTTAGCAGTCAAGATGAAACCAACCAAGAGATTTGTGTCAATGCTG 780
 Db 1037 GGTCTCTGTTATGTTAGTACCGGAGCTCGAAACCAACCAAGAGATTTGTGTCAATGCTG 1096
 QY 781 GACTTGAAGCCTGG 794
 Db 1097 GACTTGAAGCCTGG 1110

RESULT 5
ACL18332/c

ID ACL18332 standard; DNA; 636 BP.
 AC ACL18332;
 XX
 XX
 DT 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 DE DNA clone originating in barley containing SNP encoding sequence #8323.
 XX
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KW gene; ss.
 OS Hordeum vulgare; ssp. spontaneum.
 XX
 PN WO2003057877-A1.
 XX
 XX 17-JUL-2003.
 XX
 XX 16-DEC-2002; 2002MO-IB005403.
 XX
 XX 20-DEC-2001; 2001JP-00387059.
 PR
 XX 20-DEC-2001; 2001JP-00387131.
 PR
 XX 20-DEC-2001; 2001JP-00403299.
 PR
 XX 20-DEC-2001; 2001JP-00403300.
 PR
 XX 27-SEP-2002; 2002JP-00327515.
 XX
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 PA
 XX Sato K, Takeda K, Kohara Y;
 PI
 XX WPI; 2003-587127/55.
 DR
 XX
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.
 XX
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 PS
 CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-Oct-2003 to
 CC standardise OS field)
 XX
 XX Sequence 636 BP; 154 A; 179 C; 138 G; 165 T; 0 U; 0 Other;
 SQ
 Query Match 21.9%; Score 350.8; DB 8; Length 636;
 Best Local Similarity 72.1%; Pred. No. 1e-78;
 Matches 457; Conservative 0; Mismatches 177; Indels 0; Gaps 0;
 QY 373 AGAATGTTGAATGGTTGAAGCCAGCGGTTACATTTCTTCAGGAAATCTGTTCAT 432
 DB 634 AGATAGTAGTAAGTGGCTGGAGCCCTGGTGACATATTTTATCAGGAATCGGCTTCAT 575
 QY 433 CAATCTGGGGATCAAAACGCAAAAGCAATCCACCACTACCGTGAACTTAGTTCATC 492
 DB 574 CAGTCTGGTATTCAGAGGAAGTAGAACCGACGACCTACCGGAGCCGAGGTTTAC 515
 QY 493 ACTAAGGCTCTCAAGAGTGTCATTTGCAAGATGGATCTGGAACCTCTTATGAGCTCTCC 552
 DB 514 ACCAAGGTGTTCAAGGAATGCCACTCTATGACCAAGGCGGAACTTTTGAGCTTCT 455
 QY 553 CTACTTACTCTCAATGATTTGGAGCTTATGTCAAAAAGAAACCAAGCAAGATT 612
 DB 454 CTGGTAACCTCAAGTGTATGGAGCTTATGTGAAANGCAAGAACCAAGCAAGGATA 395

QY 613 AGTTGGTTGTGGCAAAAAGTTGATTCTAAGGATGATAAGGGGTTCCAGCGATTCTGGAT 672
 DB 394 TGCTGGCTATGGGAGAAGGTCAGTGTCACAGAAGACAAGGCTTTTCAGAGATTCTGGAC 335
 QY 673 ACTAGCCAGTACAGATGTAATAGCATTTCTGGCATATGAGCGTGATTTGGCCCTGGTTAT 732
 DB 334 AATGTGAGTACAAATCCACTGGATCTTTGCGCTACGAGCGCGTGTTTGGGGAGGGTTAT 275
 QY 733 GTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTGAGCTTGAAGCCT 792
 DB 274 GTTAGCACCGGTGGATTCGAGACCAAAAGGAATTTGTGGACAAGCTGGAAGCT 215
 QY 793 GGCCAGAAGGTCCTGGATGTTGGTGTGGAATTTGGTGGAGGTGACTTTTACATGCGGAG 852
 DB 214 GGCCAGAAGGTGCTTGATGTTGGTGTGTCATCGGAGGAGGCGACTTCTACATGCTGAA 155
 QY 853 ACCTTTCATGTTGAGGTTGTTGGATTTGATCTCTCGTTAATATGATTTCTTTGCCCTT 912
 DB 154 ACCTATGATGTCCATGTCTCGGCATCGATCTTTCATCAACATGGTTTCTTCGCGATC 95
 QY 913 GAGCGTTCTATTGGGCTTAAATGCTGTGTGAGTTTGGAGTAGCAGATTGCACCAAGATA 972
 DB 94 GAGCGTGCCATCGGTGCTCTCTGCTGCTTGAAGTTGAGGTTGCTGACTGCACCAACAAG 35
 QY 973 AACTACCTGATTAACCTCTTTGATGTCATCTATA 1006
 DB 34 GAATACCCAGAGAACACGTTTCGATGTCATCTACA 1

RESULT 6

ACL25710
 ID ACL25710 standard; DNA; 605 BP.
 XX
 AC ACL25710;
 XX
 DT 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 DE DNA clone originating in barley containing SNP encoding sequence #15701.
 XX
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KW gene; ss.
 XX
 OS Hordeum vulgare; var. (cul.Haruna Nijo).
 PN WO2003057877-A1.
 XX
 XX 17-JUL-2003.
 PD
 XX 16-DEC-2002; 2002MO-IB005403.
 PF
 XX 20-DEC-2001; 2001JP-00387059.
 PR
 XX 20-DEC-2001; 2001JP-00387131.
 PR
 XX 20-DEC-2001; 2001JP-00403299.
 PR
 XX 27-SEP-2002; 2002JP-00327515.
 XX
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 PA
 XX Sato K, Takeda K, Kohara Y;
 PI
 XX WPI; 2003-587127/55.

Single nucleotide polymorphism sites in barley varieties and DNA
 sequences containing them for analysis and identification of barley
 varieties and production of barley transformants with desired
 characteristics.

Disclosure: SEQ ID XX; 284pp; Japanese.

The present invention relates to oligonucleotide clones originating in
 barley (Hordeum vulgare) which contain single nucleotide polymorphisms

CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with chem and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)
 XX
 XX Sequence 605 BP; 186 A; 98 C; 151 G; 170 T; 0 U; 0 Other;

Query Match 20.3%; Score 324.6; DB 8; Length 605;
 Best Local Similarity 71.1%; Pred. No. 4.4e-72;
 Matches 429; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
 QY 150 TGGTGAATGGCTGAGAAAGCTGGCCAGGTATTTCTCTGGATTTCATTGAGAGCTGTAT 209
 Db 2 TGGAGAACTGGCTAAGACAGCTGGCATGTTCTTGCATGGATTTCATTGGAAGCTGTAT 61
 QY 210 CAAGAAGAAATGAAGTAAATCAATGGCACTACAAAATGTCAGTTTATGTGTCTGATGT 269
 Db 62 TAAAAAGATGAAGACATAATGCCATTACGAAATGCATCTTCATGTGTCTGATGT 121
 QY 270 GACTTCTCCACTCTGATTTCCACACATTCATTCGATGATATTCCTCAATTTGTT 329
 Db 122 TACATCTCCAGACCTGGTGATTCAGGACAACTCGATGATTCATATTTTCAAACTGGCT 181
 QY 330 ACTCATGTATCTTCTCATGAAGAGGTGGAAATTTGGTTGAAAGATGTTTGAATGGTT 389
 Db 182 ACTGATGTATCTTTCAGACGAGGAGGTGCGAAGCTTGTGAGAAAGATGTTAAATGGCT 241
 QY 390 GAAGCCAGGGGTTACATTTCTTCAGAGATCTTGTTCATCAATCTGGGATCAAA 449
 Db 242 CAAGTTGGTGGCATTCTCTTTAGAGATCATCTTCCATCAATCTGAGACTCAA 301
 QY 450 AGCCAAAGACATCCCAACCACTACCGTGAACCTAGTTTCTACATTAAGGCTTCAAAGA 509
 Db 302 AAGGAAAGTGAATCCGACATATATCGTGAACCAAGGTTTACACTAAGGTATTTAAGA 361
 QY 510 GTGTCATTTCAGATGATGATCTGMAACTCTTATGAGCTCTCCCTACTTACGCAATG 569
 Db 362 GGGCCATGCCATTCATCAGATGGAGCTCTCTGACTTCTCTGCTTACTTGCAGTG 421
 QY 570 TATTGGAGCTTATCTCAGAACAGAAACACAGACAGATAGTTGGTTGTGCAAAA 629
 Db 422 TGTGGAGCTTATGTGAAGAACAGAAAGATCAAAACCAAGATATGTGGCTATGGCAAAA 481
 QY 630 AGTTGATTTAAGGATGATTAAGGGGTTCCAGCGATTTCTGGATACTAGCCAGTACAAAGTG 689
 Db 482 AGTCAACTCAACAGAGATCGGGGTTTCAAGATTTTGGATATGTGCAAGTACAAAAC 541
 QY 690 TAATAGCATCTCGGATATGAGCGGTGATTTGGCCCTGGTATGTTAGCATCTGGAGGATA 749
 Db 542 CAGTGGATATACGCTATGAGCGGTGTTTGGCGAAGGTTTGTGAGCATCTGGTGGAT 701
 QY 750 TGA 752
 Db 602 TGA 604

RESULT 7
 ACL19398/c
 ID ACL19398 standard; DNA; 566 BP.
 XX
 AC ACL19398;
 XX
 XX 17-OCT-2003 (first entry)
 XX DNA clone originating in barley containing SNP encoding sequence #9398.
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

gene; ss.
 KW
 XX Hordeum vulgare.
 OS
 XX WO2003057877-A1.
 PN
 XX 17-JUL-2003.
 PD
 XX 16-DEC-2002; 2002WO-1B005403.
 PP
 XX 20-DEC-2001; 2001JP-00387059.
 PR
 XX 20-DEC-2001; 2001JP-00387131.
 PR
 XX 20-DEC-2001; 2001JP-00403299.
 PR
 XX 20-DEC-2001; 2001JP-00403300.
 PR
 XX 27-SEP-2002; 2002JP-00327515.
 XX (UYN1-) UNIV JAPAN OKAYAMA.
 PA
 XX Sato K, Takeda K, Kohara Y;
 PI
 XX WPI; 2003-587127/55.
 DR
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transfectants with desired
 PT characteristics.
 XX
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with chem and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences
 XX
 XX Sequence 566 BP; 168 A; 140 C; 97 G; 156 T; 0 U; 5 Other;

Query Match 19.0%; Score 304.2; DB 8; Length 566;
 Best Local Similarity 71.9%; Pred. No. 6.4e-67;
 Matches 407; Conservative 0; Mismatches 158; Indels 1; Gaps 1;
 QY 485 GGTTCATACCTAAGGCTTCAAAAGATGTCATTTCAGAGATGGATCTCGAAACTCTTATG 544
 Db 566 GGTTCATACCTANGTATTTAAAGNNGCCATGCCATTGATCAGNAGTGGAGCTCTCTG 507
 QY 545 AGCTCTCCCTACTTAGCTGCATAATGTATTGGAGCTTATGTCAAGAACAGAAACACAGA 604
 Db 506 AACTTCTCTCTACTTACTTGCAGGTGTGGAGCTTATGTGAAGAACAGAGATCAAA 447
 QY 605 ACCAGATTA-GTTGGTTGTGGCAAAAAGTTGATTCAGGATGATTAAGGGGTTCCAGCGA 663
 Db 446 ACCAGATATATGTGGCTATGTCGAAAGTCAACTCAACAGAGATCGGGGGTTTCAAAGA 387
 QY 664 TTTCTGATACTAGCCAGTACAAAGTGTATGATGATTCGCGATATGAGCGGTGATTTGGC 723
 Db 386 TTTTGGATATATGTCAGTACAAAACAGTGGATATTAACGCTATGAGCGGTGTTTGGG 327
 QY 724 CTTGGTTATGTAGCACTGGAGGATATGAACCAACAGAGGTTTGTCTCATGCTGGAC 783
 Db 326 CAGGTTTGTGAGCACTGTGGATTTGACATCAAAAGAAATTTTGGACTTGTGAT 267
 QY 784 TTGAAGCTGCGCAAGAGGCTCTGGAATGTTGGTGTGGAATTTGGAGGTGACTTTTAC 843
 Db 266 CTTAAACCTGGCAGAGAGTGTCTTGTATGTTGATGTTGAATCGGGGTGTTGATTTT 207
 QY 844 ATGGCGAGACCTTTGATGTTGAGTTGTTGATTTGATCTCTCCGTTAATATGATTTCC 903

Db 206 ATGCGCGAAACTATGATGTCATGTTGGCAATGATCTTTCATAAACATGTTTCA 147
 QY 904 TTTCCTTGGAGCTTCTATTTGGGCTTAAATGCTGTTGAGTTTTCAGGTAGCAGATTGC 963
 Db 146 TTTCCTTGGAGCTTCTATTTGGGCTTAAATGCTGTTGAGTTTTCAGGTAGCAGATTGC 87
 QY 964 ACCAGATATATACCTGCTGATATCTTTTGTGATCTATATAGCGTGCACACCATTCG 1023
 Db 86 ACCAGATATATACCTGCTGATATCTTTTGTGATCTATATAGCGTGCACACCATTCG 27
 QY 1024 CATATTCAGCAGAGCTGCTGCTGTTT 1049
 Db 26 CACATACAGATATACCTGCTGCTGTTT 1

RESULT 8

ID ADE82005 standard; cDNA; 487 BP.

AC ADE82005;

XX 29-JAN-2004 (first entry)

DE Arabidopsis thaliana expressed polynucleotide seq id 776.

XX genetically modified organism; transgenic organism; plant;
 KW inhibitor testing; activator testing; modifier testing; fungicide;
 KW insecticide; genetic function; genetic regulation; cellular metabolism;
 gene; 88.

XX Arabidopsis thaliana.

OS US2003115639-A1.

XX 19-JUN-2003.

XX 26-JAN-2001; 2001US-00770961.

XX 27-JAN-2000; 2000US-0178466P.

XX (GORL/) GORLACH J.

PA (ANYI/) AN Y.

PA (HAMI/) HAMILTON C M.

PA (PRIC/) PRICE J L.

PA (RAIN/) RAINES T M.

PA (YUYI/) YU Y.

PA (RAME/) RAMEAKA J G.

PA (PAGE/) PAGE A.

PA (MATH/) MATHW A V.

PA (LEDF/) LEDFORD B L.

PA (WOES/) WOESSNER J P.

PA (HAAS/) HAAS W D.

PA (GARC/) GARCIA C A.

PA (KRIC/) KRICKER M.

PA (SLAT/) SLATER T.

PA (DAVI/) DAVIS K R.

PA (ALLE/) ALLEN K.

PA (HOFF/) HOFFMAN N.

PA (HURB/) HURBAN P.

XX Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;

PI Ramakia JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;

PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;

PI Hurban P;

XX WPI; 2003-810930/76.

XX Novel Arabidopsis thaliana nucleic acids useful for generating
 PT genetically modified transgenic organisms, for screening biologically
 PT active agents such as fungicides, insecticides.

XX Claim 1: SEQ ID NO 776; 44pp; English.

XX

CC The invention describes a nucleic acid (1) comprising a sequence capable
 CC of hybridising under stringent conditions to any one of 999 fully defined
 CC Arabidopsis thaliana sequences (1) as given in specification e.g., 360,
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (1) is useful as a
 CC hybridisation probe to complementary molecules in a cDNA library. (1) is
 CC also useful for generating genetically modified and transgenic organisms,
 CC usually plant cells and plants. A protein encoded by (1) is useful in
 CC screening assays to determine the effect of candidate inhibitors,
 CC activators or modifiers of the gene product. The protein is also useful
 CC for screening biologically active agents e.g., fungicides and
 CC insecticides. A genetically modified cell, comprising an exogenous
 CC nucleic acid, where the nucleic acid comprises transcription regulatory
 CC sequences operably linked to a sequence capable of hybridising under
 CC stringent conditions to (1) is useful in the study of genetic function
 CC and regulation, for alteration of the cellular metabolism and for
 CC screening compounds that may affect the biological function of the gene
 CC or gene product. This sequence represents an Arabidopsis thaliana
 CC polynucleotide of the invention.

XX SQ Sequence 487 BP; 129 A; 88 C; 131 G; 139 T; 0 U; 0 Other;

Query Match 19.0%; Score 303.6; DB 9; Length 487;

Best Local Similarity 76.5%; Pred. No. 8.6e-67;

Matches 372; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 543 TGAGCTCTCCCTACTTAGCTGCAAAATGTTTGGAGCTTATGTGCAAAACCAAAACCA 602

Db 1 TGAGCTCTCTATGTTGGCTGCAAAATGTTTGGAGCTTATGTGCAAAACCAAAACCA 60

QY 603 GAACGAGATTAGTTGGTTGGCAAAAGTTGATTTCTAAGGATGATAGGGGTTCCAGCG 662

Db 61 GAATCAGATTGCTGGGATATGGCAAAAGTCAGCGTGGAGATGCAAGGATTTCCAGCG 120

QY 663 ATTTCTGATAGTACCCAGTACAGTGTATATAGCATTTCTGCGATATGACGCTGATTTGG 722

Db 121 TTTCTTGGACATGTTTCAATACAGTCTAGTGGGATCTTGGCTATGAGCGTGTCTTTGG 180

QY 723 CCTGGTTATGTTAGCACTGGAGGATATGAAACCAACCAAGATTTGTGTAATCTGGA 782

Db 181 GGAGGATATGTTAGCACTGGTGGATTTGAGACACTAAAGAAATTTGTGGCGAAGATGA 240

QY 783 CTTGAAGCTTGGCCAGAGGTCCTGGATTTGGTGTGGAAATTTGGTGGAGGTGACTTTTA 842

Db 241 CCTTAAACCGGACAGAAAGTCTTAGATTTGGTGTGGTATCGGTGGAGGTGACTTTTA 300

QY 843 CATGGCGGAGACCTTTGATTTGAGGTTTGGATTTGATCTCTCGTTTAAATATGATTTC 902

Db 301 CATGGCTGAGAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 360

QY 903 CTTTGGCTTGGAGCTTCTATTTGGGCTTAAATGCTGTTGAGTTTGGAGTGGAGATG 962

Db 361 TTTGCACTGGAGCGGGCCATTGGACTCAAATGCTCAGTTCGAGTTTGAAGTCCGCTGATTG 420

QY 963 CACCAAGATAAACTACCCCTGATTAACCTTTTGTATGTCATCTATAGCCGTCAGCACCATTCT 1022

Db 421 CACCAACCAAAACATATCCCGATTAATTCCTTTGATGTCATTTTACAGCGGTGACATTTCT 480

QY 1023 GCATAT 1028

Db 481 GCATAT 486

RESULT 9

ACL25711

ID ACL25711 standard; DNA; 611 BP.

XX ACL25711;

XX 27-OCT-2003 (revised)

DT 17-QCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #15702.

XX

Matches 403; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 846 GCGGAGACCTTTGAGTTCAGGTTGATGATTCCTCCGTTAATATGATTCCTT 905

Db 777 GNCGAAACATATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 718

QY 906 TGCCCTTGAGGCTTATGCTGCTTAAATGCTGTTGATGTTGATGTTGATGTTGAT 965

Db 717 TGCACTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658

QY 966 CAGATAAATACCTGATTAATCTTTGATGCTATGCTGCTGCTGCTGCTGCTGCTGCT 1025

Db 657 CAGGAGACATACCAACATATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 598

QY 1026 TATTGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085

Db 597 CATACAGATAAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538

QY 1086 AGTTCTAATCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145

Db 537 GGTCTAATCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478

QY 1146 CATTAAGCAGAGGCTATGATCTCCATGATGTAAGGATGTAAGGATGTAAGGAT 1205

Db 477 CATTAAGCAGAGGCTATGATCTCCATGATGTAAGGATGTAAGGATGTAAGGAT 418

QY 1206 TGCTGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265

Db 417 TGCTGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 358

QY 1266 GAAGGACTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1325

Db 357 GAGGAGCTAGCCGAGTGTGAAAGAAACAAAGATGATTTCTAGCGACTTTGGTCAGGA 298

QY 1326 GGAATTAATGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385

Db 297 GCACTATGACATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238

QY 1386 GCAAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421

Db 237 GCAGAGTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202

RESULT 11

ACLI8347/c

ID ACLI8347 standard; DNA; 575 BP.

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

XX AC ACLI8347;

PI Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA sequences containing them for analysis and identification of barley varieties and production of barley transformants with desired characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in barley (Hordeum vulgare) which contain single nucleotide polymorphisms (SNP). The oligonucleotides may be used for analysis of SNPs among barley varieties, identification of particular varieties and genotype-phenotype analysis, isolation of specific genes and creation of new varieties by transformation of barley varieties with them and production of new barley varieties with desired properties. The present sequence represents an oligonucleotide clone DNA sequence featured in the specification. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published-pet-sequences. (Updated on 27-OCT-2003 to standardise OS field)

SQ Sequence 575 BP; 128 A; 160 C; 127 G; 159 T; 0 U; 1 Other;

Query Match 17.5%; Score 280.6; DB 8; Length 575;

Best Local Similarity 68.9%; Pred. No. 6.3e-61;

Matches 385; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 237 CTACAAAATGTCAGTTTATGTTGCTGATGTTGCTTCTCCACTCTCCAGTTTCCCACC 296

Db 559 CCACAGCAACATACCTTTCATGTCGCCCGACGTCGCGGAGCTCAAGATCAGGA 500

QY 297 ACATTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 356

Db 499 CAACTCCATGACATGCTCTTCTCAACTGGCTGCTCATGTACCTCAACGACGAGAGGT 440

QY 357 GGAATAATTTGTTGGAAGAATGTTGAATGTTGGAAGCTGAGCGGGGTTACNTTTCTTCAG 416

Db 439 TGAGAGCTGATTTGGCAGAAATAGTGAAGTGGCTGAGGCTGCTGGGACATATTTTATCAG 380

QY 417 AGAATCTTTGTTCCATCAATCTGGGATCAAAACGCAAAAGCAATCCACCACTACCG 476

Db 379 GGAATCGTGTCTCCATCAGTCTGCTGATTCAGAGGAAAGTGAACCCGACACTACCG 320

QY 477 TGAACCTAGTCTTACACTAAGGCTTCAAGAGTGTCTATTTGCAAGATGATCTGGAAA 536

Db 319 GGAGCGAGGTTTACACCAAGGTTTCAAGGATGCTCTCTATGACCAAGAGGGA 260

QY 537 CTCTTATGAGCTCTCCCTACTTACCTGCAATGTTATGAGCTTATGTCAGAACAGAA 596

Db 259 CTCTTTTGAAGCTTCTCTGTTAACTTCCAAAGTGTATTTGAGCTTATGTAAGAACAGAA 200

QY 597 AAACCCAGAACCAATGATGTTGTTGGCAAAAGTGTATTTAAAGGATGATTAAGGGGT 656

Db 199 GAACCCAGAACCAATGATGTTGTTGGCAAAAGTGTATTTAAAGGATGATTAAGGGGT 140

QY 657 CCAGGATTTCTGGAATACTAGCCAGTACAGTGTATATGATGATTTCTGCGATATGAGCTGT 716

Db 139 TCAGAGATTTCTGGAATACTAGCCAGTACAGTGTATATGATGATTTCTGCGATATGAGCTGT 80

QY 717 ATTTGGCCCTGGTTATGTTAGCTCTGGAGATATGAACCAACCAAGAGTTTGTGTCAAT 776

Db 79 GTTTGGGAGGGTTATGTTAGCACCCTGGTGGATTCGAGACCCACAAAGGAATTTGTGGACA 20

QY 777 GCTGGACTTGAAGCTTGGC 795

Db 19 GCTGGACTTGAAGCTTGGC 1

RESULT 12
ACLI8333/c

CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)

XX SQ Sequence 622 BP; 118 A; 187 C; 173 G; 144 T; 0 U; 0 Other;

Query Match 17.0%; Score 271.8; DB 8; Length 622;

Best Local Similarity 69.1%; Pred. No. 1.1e-58;

Matches 386; Conservative 0; Mismatches 172; Indels 1; Gaps 1;

QY 858 TGATGTCAGGTCCTGGATTCGATCTCCGTTATATGATTCCTTTGCCCTTGAGCG 917
 DB |||||
 QY 620 TGATGTCATGTCCTGGCATCGATCTTCC-TCACATGGTTTCTTCGGATCGAGCG 562
 DB |||||
 QY 918 TTTATATGGGCTTAAATGTCGTGTTGAGTTGAGGTAGCAGATTGACACAGATAAATA 977
 DB |||||
 QY 561 TGCCATCGGTCCTCGCTCGGTGAGTTGAGTTGCTGACTGCACCCACCAAGGAATA 502
 DB |||||
 QY 978 CCTGTGATACCTCTTTGATGTCATCTATAGCCTGACACCATTTCTGCTATTCAGGACAA 1037
 DB |||||
 QY 501 CCCAGAGAACACGTTGATGTCATCTACAGCCGTGACACCATCTCCACATTCAGACAA 442
 DB |||||
 QY 1038 GCTCGGTTGTTAGTCTCTTACAAATGTTGAGCCAGCAGGTAAAGTTCTAATCAG 1097
 DB |||||
 QY 441 GCTCTCTGTTGAGAACTCTTCAAGTGGCTGAAAGCTGTGGCAAGTCTGATCAG 382
 DB |||||
 QY 1098 TGATTAATGCAAGAAAGCTGTTCCACCTCACTGAAATTCGGCGTTACATTAAGCAGAG 1157
 DB |||||
 QY 381 TGACTACTGCAAGAGCCCTGGAGCCCATCAGAGGAATTCGCTGGGTACATCAAGCAGAG 322
 DB |||||
 QY 1158 GGGATATGATCTCATGATGTAAGAGGAATATGCGGAGATGCTTAAAGATGCTGGATTGT 1217
 DB |||||
 QY 321 AGGCTATGACCTCCATGATGTAAGAGCTTACGAAAGATGCTTGAGATGCGGTTCCA 262
 DB |||||
 QY 1218 TGATGTTCTTCCGAGGATGAGTACGAGGATTCATTCGAGTTCTACGGAGGAACTAGA 1277
 DB |||||
 QY 261 CGACGTCTATCCGCGAGACCGACCCAGATTCCTTGAGGGTCTTGAGAGAGGAGCTGG 202
 DB |||||
 QY 1278 GACTGTGAGAGGAAAGAGATGTTTCATTAGTATGATTTCTCTGAGGAGGATTAACAATGA 1337
 DB |||||
 QY 201 CGAGACCGAGAGAACAGAGAGGCTTCTGCGGAGCTTCAGCCAGGAGACTACGAGCA 142
 DB |||||
 QY 1338 CATTTGTTGGGTTGGATGATGATGTTGCGAGGACTGCGAGGTTGAGCAAGATGGG 1397
 DB |||||
 QY 141 CATCTGCAACCGGTTGGAGCGGAGCTGAAGCGGAGCTCCGCGGAGGAGTGGGG 82
 DB |||||
 QY 1398 TCTGTTCTTCCCAAGAG 1416
 DB |||||
 QY 81 GCTGTTTCATCGGACCAAG 63
 DB |||||

RESULT 14

ACL25708

ID ACL25708 standard; DNA; 589 BP.

XX AC ACL25708;

XX AC ACL25708;

XX AC ACL25708 (revised)

XX AC ACL25708 (first entry)

XX AC ACL25708 (first entry)

XX AC DNA clone originating in barley containing SNP encoding sequence #15699.

XX AC Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 XX AC gene; ss.

XX AC Hordeum vulgare; var. (cul.Haruna NiJo).

XX PN WO2003057877-A1.
 XX PD 17-JUL-2003.
 XX PF 16-DEC-2002; 2002WO-IB005403.
 XX PR 20-DEC-2001; 2001JP-00387059.
 XX PR 20-DEC-2001; 2001JP-00387131.
 XX PR 20-DEC-2001; 2001JP-00403299.
 XX PR 20-DEC-2001; 2001JP-00403300.
 XX PR 27-SEP-2002; 2002JP-00327515.
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 XX PA Sato K, Takeda K, Kohara Y;
 XX PI WPI; 2003-587127/55.
 XX DR Single nucleotide polymorphism sites in barley varieties and DNA
 XX PT sequences containing them for analysis and identification of barley
 XX PT varieties and production of barley transformants with desired
 XX PT characteristics.
 XX PS Disclosure; SEQ ID XX; 284pp; Japanese.
 XX CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)

XX SQ Sequence 589 BP; 161 A; 120 C; 166 G; 142 T; 0 U; 0 Other;

Query Match 16.8%; Score 268.6; DB 8; Length 589;

Best Local Similarity 71.1%; Pred. No. 7e-58;

Matches 355; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1 CACACCGTTGATTTAAACCATTTGAAGCTATGATGCTGATTCCTCAAGCTTCTGATCTTGAC 60
 DB |||||
 QY 61 AAAGAGAAAGCTCTGAGATTCCTTCAATGCTTCGCTCTTGAAGGAAATGCTCTTG 120
 DB |||||
 QY 149 AAGGAGAGCGCCCGGAGATGATGCTTCTTACTTCTTCAATGAGGAAATCAGTCTG 208
 DB |||||
 QY 121 GAACCTGGGCTGATTTGCTGCTTTTACTGCTGAAATGCTGAGAAAGCTGGCCAGTT 180
 DB |||||
 QY 209 GAGCTTGGCGGGGATAGGTCGCTTTACTGAGAGAACTGGCTAAGACAGCTGGGCTGTT 268
 DB |||||
 QY 181 ATTGCTCTGGAATTTCTTGAAGAGTGTCTATCAAGAGAAATGAATGAATCAATGGGCACTAC 240
 DB |||||
 QY 269 CTTCGATGATTTCTTGAAGATGATGATTAAGAGATGAAGCAATTAATGGCCATTAC 328
 DB |||||
 QY 241 AAAAATGCAAGTTTATGCTGCTGATGATGATCTTCCCACTCTGATTTCCCAACAT 300
 DB |||||
 QY 329 GAAATGATCATCTTTCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 388
 DB |||||
 QY 301 TCATTGGATGATTAATTTCTCCAAATTTGTTTACTCATGATGATGATGATGATGATGATGATGAT 360
 DB |||||
 QY 389 TCGATTGATGATTAATTTTCAAACTGGCTACTGATGATGATGATGATGATGATGATGATGATGAT 448
 DB |||||
 QY 361 AATTGGTTGAAAGAAATGTTGAAGTGTGAAGCCAGGGGTTTACATTTCTTCAGAGAA 420
 DB |||||
 QY 449 AAGCTTGAAGAAAGATGTTAATGGCTCAAGTTGGTGGCCATATCTTCTTTAGAGNA 508
 DB |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2004, 00:19:09 ; Search time 3581 Seconds
(without alignments)
3944.376 Million cell updates/sec

Title: US-10-031-331B-40
Perfect score: 2515
Sequence: 1 HTVDLTIEMQLDSQASLD.....KLRTAKGEQWGLFVAKKX 473

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xlh
-Q/cg2_1/USPTO.spool/US10031331/runat.26072004.120351.14986/app.query.fasta_1.647
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=diodes62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031331@CGN.1.1.2810 @runat.26072004.120351.14986 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_vrt.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_tod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1937	77.0	2333	11	AY103779	Zea mays
2	1420	56.5	1012	11	AY105261	Zea mays
3	1413	56.2	1211	10	BH574969	E07 Triph
4	1203	47.8	1022	14	CK206655	FGAS01825
5	1183	47.0	713	10	BE231445	S
6	1162	46.2	767	12	BM408047	BM408047 EST582374
C 7	1158.5	46.1	1115	14	CK166076	CK166076 FGAS05012
8	1111	44.2	869	14	CF513471	CF513471 CABud0007
9	1106	44.0	781	14	CF208434	CF208434 CAB20003
10	1092	43.4	1127	10	BE420987	BE420987 HMM004.F0
11	1066	42.4	822	10	BE040284	BE040284 OE03A05.O
C 12	1065.5	42.4	733	13	BQ965516	BQ965516 QHB22C02
13	1057	42.0	706	13	BQ120291	BQ120291 EST605867
14	1055	41.9	805	10	BE040460	BE040460 OE06G12.O
15	1038	41.3	721	14	CF482195	CF482195 POL1.5.B0
16	1029	40.9	941	14	CK265554	CK265554 EST711632
17	1027	40.8	770	14	CD320161	CD320161 G608.116A
18	1025	40.8	753	14	CB970167	CB970167 CAB10003
19	1024	40.7	718	14	CA174643	CA174643 SCJFS101
20	1024	40.7	720	14	CF508254	CF508254 USDA-PP.1
21	1018	40.5	731	13	BQ855322	BQ855322 QGB2SN08
22	1012.5	40.3	772	9	AJ611652	AJ611652 QGB11652
23	1010	40.2	736	13	BQ852610	BQ852610 QGB18H22
24	1005	40.0	750	12	BM406709	BM406709 EST581036
25	1004	39.9	745	13	BQ762263	BQ762263 EB-co1.SQ
26	1002	39.8	682	13	BQ856500	BQ856500 QGB4H12.Y
27	998	39.7	700	14	CA289955	CA289955 SCAGP1801
28	991	39.4	806	14	CK195203	CK195203 FGAS00364
29	984	39.1	737	14	CF133358	CF133358 WHB4357.A
30	969	38.5	689	13	BQ119788	BQ119788 EST605367
31	961	38.2	672	14	CD232049	CD232049 SSI.31.D1
32	957	38.1	852	12	BG343052	BG343052 HVSMEG000
33	950	37.8	621	13	BU635947	BU635947 043C05.IN
34	950	37.8	788	14	CA295025	CA295025 SCQGLV101
C 35	948	37.7	848	14	CF513552	CF513552 CABud0007
36	943	37.5	689	12	BG907548	BG907548 Tair1160H
37	942	37.4	731	13	CA100170	CA100170 SCVPCLE604
38	941	37.4	802	12	BJ573651	BJ573651 BJS73651
C 39	940	37.3	601	12	BM302134	BM302134 MCA047F06
40	938	37.3	654	12	BJ280399	BJ280399 BJS73651
41	938	37.3	771	14	CA277433	CA277433 SCACSD201
42	938	37.3	853	14	CK204209	CK204209 FGAS01274
43	938	37.3	684	14	CA216364	CA216364 SCRLFL402
44	937	37.3	684	14	CA216364	CA216364 SCRLFL402
C 45	933	37.1	720	12	BJS73369	BJS73369 BJS73369

ALIGNMENTS

RESULT 1
AY103779
LOCUS AY103779
DEFINITION Zea mays P00126548 mRNA sequence.
ACCESSION AY103779
VERSION AY103779.1 GI:21206857
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2333)

AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S.,
 Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2333)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.

FEATURES
 Location/Qualifiers
 1..2333
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:4577"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed DuPont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 4,066-2331 Length: 2333
 Score: 1937.00 Matches: 356
 Percent Similarity: 87.08% Conservative: 55
 Best Local Similarity: 75.42% Mismatches: 61
 Query Match: 77.02% Indels: 0
 DB: 11 Gaps: 0

US-10-031-331b-40 (1-473) x AY103779 (1-2333)

QY 1 HistValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 Db 488 CACTCCAGTGGCTCACTGTCGAGTCCATGATGCTCGACTCCGCGCGCGACCTCGAC 547
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu 40
 Db 548 AAGGAGAGGACCGACGATCTCTCTGCTTCCTCTTACAGAGGAGATCAGTCTTA 607
 QY 41 GluLeuGluAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60
 Db 608 GAATCGGTGGTGGAAATGGACGCTTTACTGGAGATCTGGCAAAAGAGCTGGGACGTT 667
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
 Db 668 CTGGCCTAGACTTTATTGAAGTGTGCTTTAAGAGAGACCATTAATGGGATCATC 727
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPhe 100
 Db 728 AAGACATACCTTCAGGTGGCGGATGATGATCTTAACGACTTGAAGATTGAAGATTAAC 787
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120
 Db 788 TCTGTGTGATCTGATATTTTCAAACTGGCTTAATGATATCTTTCAGATGAGGAGTCCA 847
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePheArgGlu 140
 Db 848 AAGCTGTGGGGAATGCTTAATGATTAAGTGGAGGACCATATTCTTCTTGAAGA 907
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160

Db 908 TCATGTTTTCCACCAATCTGGAGATTCCTAAAGGAAGGTGAACCAACACACATATCGAGAA 967
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 Db 968 CCAAGGTTTTATACCAAGGTATTTAAAGAGGGCCATCTTTGATCAAGATGAGGTTGG 1027
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200
 Db 1028 TTGTAACCTTCTCTAGTGACCTGTAAATGCAATGGGGCTTATGTCAAAACCAAGAGAA 1087
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysGlyPheGln 220
 Db 1088 CAAACCCAGATATGCTGTTATGGAAGAGGTAAATCAACAGACACAGAGATTTTCAA 1147
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
 Db 1148 AGATCTCTGGACACCTGCAATACAAACAGAGTGGATATTACGTTATGAGCGTGTCTTT 1207
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 Db 1208 GGTGAAGGTTTTGTGAGCCTGTTGGATCGAGACTACAAAGGAATTTGTGGGCGATGCTC 1267
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysPhe 280
 Db 1268 GATCTTAAACCGGCGCAGAAAGTACTTGTGATGTCGATGTGAATTCGAGCGCGGCGCTT 1327
 QY 281 TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300
 Db 1328 TACATGGCTGCAACTATGATGATGCTCATGTTCTTGGTATTGATCTTTGCGGTGAACATGTT 1387
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 Db 1388 TCATTTGCAATGAACGTGCCATTCGACGCAAGTGTCTGTGTAATTCGAATTCGAT 1447
 QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgPheThrIle 340
 Db 1448 TGCACCAACAGAGGATTACCCAGAAATAGTTTGAAGTCACTACAGCCGTGACACCATC 1507
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
 Db 1508 CTTTCATACAGACNAAGCTGCTCTGTTCAAGAGCTTCTTCAATGGGTGAAGCCGCGC 1567
 QY 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
 Db 1568 GGCAGAGTCTTAATCAGCGACTACTGTGAAGATCTTGAAGAACCATCAGAGATTTGCT 1627
 QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
 Db 1628 GCGTACATTAAGCAGAGAGGCTATGACCTTCACAGCTGAGGCTTATGACAGATGCTG 1687
 QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
 Db 1688 AAGGATGCTGGTTTTTCATAATGTCATCGGAAGATCGCACTGACGAGTTCTTGAATGT 1747
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
 Db 1748 CTACAGAGGAGATAGGTGAGTTGAAGAGAACCAAGAGACGCTTCTCGGACAGCTTACC 1807
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTrpAsnAspLysLeuArgThrAlaLys 460
 Db 1808 CAGGAGGATATGACGACATTTGTAATGGCTGGACCGAGCTGAACGAGCTCTGCC 1867
 QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
 Db 1868 GCGGAGCAGAGGTGGGGTGTGTTTCATTGCCACCAAG 1903

RESULT 2
 AY105261
 LOCUS Zea.mays.PC0129777 mRNA sequence. 1012 bp mRNA linear HTC 16-OCT-2002
 DEFINITION AY105261
 ACCESSION AY105261.1 GI:21208339
 VERSION AY105261.1
 KEYWORDS HTC.

SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL
Unpublished (2002)
REFERENCE
AUTHORS
Coe, E.H.
TITLE
Direct Submission
JOURNAL
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1. 1012
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637558"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Alignment Scores:
Prod. No.: 8,99e-167 Length: 1012
Score: 1420.00 Matches: 262
Percent Similarity: 85.80% Conservatives: 28
Best Local Similarity: 77.51% Mismatches: 47
Query Match: 56.46% Indels: 1
DB: 11 Gaps: 0

US-10-031-331B-40 (1-473) x AY105261 (1-1012)

QY 104 ValillePheSerAenTtpLeuLeuMetTyrLeuSerAspGluValGluAenLeuVal 123
DB 1 CIGATATTTTCARATTTGGTATTGATGATCTTTTCAGATGAAGAGGTTGACAGCTACTT 60
QY 124 GluArgMetLeuLysTtpLeuLysProGlyGlyTyrIlePhePheArgGluSerCysPhe 143
DB 61 CAGAGAATGGTAAATGGTTGAAGGTTGGTGGCTATATCTCTTTAGGGAATCTTGCTTC 120
QY 144 HisGlnSerGlyAspHisLeuArgLysSerAenProThrHisTyrArgGluProArgPhe 163
DB 121 CATCAATCTGGAGATTCARAAAGGAAGTAAATCCGACACACTATATAGGCAACCAAGTTT 180
QY 164 TyrThrLysAlaPheLysGlyCysHisLeuGlnAspGlySerCysLysSerTyrGluLeu 183
DB 181 TATACTAAGGTTTTCAGAGATGCAATCTTGTATCAAGATCGGAATCTTTCGACTT 240
QY 184 SerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAenLysLysAenGlnAenGln 203
DB 241 TCTCTGGTTACTTGCAGGTATTGGTCTTATGTATTAACCAAGAAAAACCAACACAG 300
QY 204 IleSerTtpLeuTyrGlnLysValAspSerLysSerLysGlyPheGlnArgPheLeu 223
DB 301 ATATGTGGCTATGGCAAAAGTCCATCTTACAGACATGAAGATTTCAAGATTTTTC 360

QY 224 AspThrSerGlnTyrLysCysAenSerIleLeuArgTyrGluArgValPheGlyProGln 243
DB 361 GCAATGTGCTGATCAAAAGCCAGTGGGANTATTAGTTACGAGCGCATTTTGGAGAGG 420
QY 243 TyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeuAspLeuLys 263
DB 421 TTATGTGAGCACTGGTGGAGTTGAGACATACAAAGATTTGTGGACAGCTGGATCTCAA 480
QY 263 sProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyGlyAspPheTyrMetAl 283
DB 481 ACCTGGACATAGCTGCTTGTATGTTGGATGGGAATGGGGAGGTGACTTTTATATGGC 540
QY 283 aGluThrPheAspValGluValGlyPheAspLeuSerValAenMetIleSerPheAl 303
DB 541 CGAAATATATGATGCTATGTTGTTGTTGATGATCTTCCATTAACATGGTATCATTTGC 600
QY 303 aLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAspCysThrLys 323
DB 601 ACTTGACGCTGCCATTTGGGGCGAGTTGCTCAGTGGAGTTTGAAGTTGCTGATTGCATAC 660
QY 323 sIleAsnTyrProAspAenSerPheAspValIleTyrSerArgAspThrIleLeuHisI 343
DB 661 GAAGACATACCCAGACACACATTTTGATGTCATATACAGCCGCGTACTATCTTCCAT 720
QY 343 eGlnAspLysProAlaLeuPheArgSerPheTyrLysTtpLeuLysProGlyGlyLysVa 363
DB 721 ACATGACAAACCTCTTGTTCARAAAGTCTTTCAGTGGCTGAAAGCTGGGGCAAGGT 780
QY 363 lLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAlaLalaTyrIl 383
DB 781 CCTTATCAGTGACTACTCTGGAGAGTCTCTGGAAACCATCAGAGGAATTTGCGAGGTACAT 840
QY 383 eLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAl 403
DB 841 TATGACAGAGGTTATGACCTACATCTCTGGAGGCTTATGACAGAGTGTGAGAGGTGC 900
QY 403 aGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLys 423
DB 901 TGGTTTTCGTGATGTCATCTGAGGATCGAATGATGATGCTTCTGGTGTGTTTATAGATA 960
QY 423 sGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
DB 961 GGAGTTAGCTGAATTTGAAAGAAACAGACAGATTTCTCTGCTGACTTCACC 1012

RESULT 3

BE574969 1211 bp mRNA linear EST 07-DEC-2001
E07 Triphysaria versicolor root-tip, early DMBQ-induced transcript
cDNA library Triphysaria versicolor cDNA, mRNA sequence.

LOCUS

BE574969

ACCESSION

BE574969

VERSION

BE574969.1

KEYWORDS

EST

SOURCE

Triphysaria versicolor

ORGANISM

Triphysaria versicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Lamiales; Orobanchaceae; Rhinanthaceae;
Triphysaria.

REFERENCE

1 (bases 1 to 1211)

AUTHORS

Matvienko, M., Torres, M.J. and Yoder, J.I.

TITLE

Transcriptional responses in the hemiparasitic plant Triphysaria

JOURNAL

Plant Physiol. 127 (1), 272-282 (2001)

MEDLINE

21437952

PUBMED

1155755

COMMENT

Contact: Yoder, J.I.
John I. Yoder Research Lab, Dept. of Vegetable Crops
University of California at Davis
137 Asmunden Hall, One Shields Drive, Davis, CA 95616, USA
Tel: 530 752-1741
Fax: 530 752 9659
Email: jyoder@ucdavis.edu
length = 1211 bp.

FEATURES

Location/Qualifiers
 1. .1211
 /organism="Triphysaria versicolor"
 /mol_type="mRNA"
 /db_xref="taxon:64093"
 /tissue_type="root-tips -5mm length"
 /dev_stage="3-4 weeks growth"
 /lab_hosts="E. coli"
 /clone_lib="Triphysaria versicolor root-tip, early
 DMBQ-induced transcript cDNA library"
 /note=Vector: pCR2.1 TA Cloning System, Invitrogen,
 Carlsbad, CA; PCR-based suppression subtractive
 hybridization cDNA library"

ORIGIN

Alignment Scores:
 Pred. No.: 9,22e-166 Length: 1211
 Score: 1413.00 Matches: 260
 Percent Similarity: 90.25% Conservative: 27
 Best Local Similarity: 81.76% Mismatches: 31
 Query Match: 56.18% Indels: 0
 DB: 1.0 Gaps: 0

US-10-031-331B-40 (1-473) x BE574969 (1-1211)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 DB 258 CATTCTGTGACCTGACCTGGAATCGATCGATGCTTGTATCCAAAGCATCTGATCTCGAT 317
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40
 DB 318 AAGGAAGAAGCCCTGAGGTGTATCTATCTCTCCACCATATGGAAGGCAAGTCGCTCTT 377
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
 DB 378 GAACTGGGAGCTGTATGTGCGCTTTCACCTGGTGAATTAGCTAAGAGGCCCAAGAGCTT 437
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLeuAsnGluValIleAsnGlyHisTyr 80
 DB 438 GTAGCTTTAGACTTTCGGAAGCATATTAAGAGAGATGAAGCCCTCAATGCCGATCAT 497
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLysSerPheProHis 100
 DB 498 AAAATGTCAAATTCATTTGTGCTGATGTCACATCCCAAGCTTGAACCTTCCTGAAGGA 557
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120
 DB 558 TCACCTGACGTGATATTCGAACTGCTATGATGATCTTTCAGACAATGAGGTGGAG 617
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
 DB 618 AATCTTGCAGAGAGATGGTCAATGTTGATGATGTTGGTGGCTATATATTTTCAGAGAA 677
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
 DB 678 TCATGTTTCCATCATGTCGCTGACCAACAAAGAGAGATATATCCGACTCATATTCGCGAA 737
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 DB 738 CCAAGATTTATACCAAGTTGTTCAAGATGCCATATGTCATGATGATCTGGAATTCATCA 797
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysValIleGlyAlaTyrValArgAsnLysLysAsn 200
 DB 798 TTGACCTGCTCTTATGTTGTTAAGTGCATCGGAGCTTATGTGAAAAACAAAAAGAAAT 857
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
 DB 858 CAGATCAGATCTGCTGGATTTGGCAAAAGGTTAGCTCAGATCATGACAGAGGATTCAG 917
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
 DB 918 AAATCTTGGATATGTCGAATATTAATGTATGTCATATGCGTATGACGAGATTTTC 977

QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 DB 978 GCACAGGCGATATGTGACGACCGGTGGACTAGAACAACTAAAGATTTGACCAACTG 1037
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
 DB 1038 GACCTTAAGTCGGCCAGGAAGTCTAGACGTGGGTGGCATTCGAGAGGATGATTTTC 1097
 QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
 DB 1098 TACATGGCTGACGAGCATGATGTTTCATGTTGTTGGCATTCGACCTCTCGGTCAATGATT 1157
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluVal 318
 DB 1158 TCTTTTGTCTCGAGCGTGCATGCTGCTCAGTGTGCTGTGAGTTGAGGTT 1211

RESULT 4

CK206655

LOCUS

DEFINITION

aestivum cDNA, mRNA sequence.

ACCESSION

CK206655

VERSION

CK206655.1

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1022 bp mRNA linear EST 08-DEC-2003
 FGAS018258 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
 aestivum cDNA, mRNA sequence.

CK206655.1 GI:39569045

EST.

Triticum aestivum (bread wheat)

Triticum aestivum

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.

1 (bases 1 to 1022)

Allard, F., Crosby, W.L., Danyluk, J., Rudes, F., Frick, M., Gaudet, D.,
 Links, M.G., McCarthy, E.L., Gullik, P., Hrycan, L.D., Larocque, A.,
 Penniket, C., Roach, J.L. and Sarhan, F., Muzak, I., Nilson, D.,
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops
 Unpublished (2003)

Contact: Wm L Crosby

Bioinformatics

University of Saskatchewan, Department of Computer Science

1C101 Engineering Building, 57 Campus Drive, Saskatoon,
 Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas.este@usask.ca

This sequence is the direct result of the Base calling software
 Phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [9,841].

Plate: L5B008 row: C column: 07.

Location/Qualifiers

1. .1022

/organism="Triticum aestivum"

/mol_type="mRNA"

/db_xref="taxon:4565"

/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"

/notes=Vector: pCMV.SPK16; Crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methyated dCTP thereby protecting from internal cleavage
 with NotI."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 1,68e-139 Length: 1022
 Score: 1205.00 Matches: 217
 Percent Similarity: 85.91% Conservative: 39
 Best Local Similarity: 72.82% Mismatches: 42
 Query Match: 47.83% Indels: 0
 DB: 14 Gaps: 0

US-10-031-331B-40 (1-473) x CK206555 (1-1022)

QY 175 AspGlySerGlyValSerThrGluLeuSerLeuLeuSerCysValLeuGlyValAlaTyr 194
 DB 27 GACCAAGATGGGAATCTCTTCTGAGCTTCTGCGCACTTCCAGAGTCAGTGGAGCTTAT 86
 QY 195 ValArganLysLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 214
 DB 87 GTGAAGAGCAAGAGAACACAGAACAGATATGTTGGCTATGGGAAGAGGTCAGTGCGACA 146
 QY 215 AspAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeu 234
 DB 147 GAAGCAAAAGGCTTTCAGAGATTCCTGGCAATGTGCAGTACAAATCCACCTGGGAATCTTG 206
 QY 235 ArgTyrGluArgValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLys 254
 DB 207 CATTATGAGCGCGTGTGGGAGGGTATGTGACGACCGGTGGATTCGAGCACCAAAA 266
 QY 255 GluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGly 274
 DB 267 GAATTTGTGGCAAGCTGGATCTGAAAGCTGGCGAGAGGTCTCGATGTTGGGTGGCGGC 326
 QY 275 IleGlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAsp 294
 DB 327 ATCGAGGAGGTGACTTCTACATGGCTGAACTATGATGTCATGCTCTCGGCATCGAT 386
 QY 295 LeuSerValAnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaVal 314
 DB 387 CTTTCCATCACATGGTTTCTGCGGATCGAGCGTCCATCGGCGCTCGTGTCTGGTT 446
 QY 315 GluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIle 334
 DB 447 GAGTTTGAGTGTGTGATGTCACCAAGCAATATCGCAGAGAACACAGTTTGTGATGTGATC 506
 QY 335 TyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyr 354
 DB 507 TACAGCGTGACACCTCTCCATTCAGACAAACCTGCTCTGTTTCAAGAACTCTCTTC 566
 QY 355 LysTyrLeuLysProGlyGlyValLeuLysValLeuLysAspTyrCysLysLysAlaGlyPro 374
 DB 567 AAGTGGCTCAAGCGCTGGGGGCAAGTCTGTATCAGTCACTGATCGAGGAGCGCTGGGACA 626
 QY 375 ProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLys 394
 DB 627 CGGTCAAGAGATTTGCTGCATACATCAAGCAGAGAGGCTATGACCTCCATCAGCTGAAG 686
 QY 395 GluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThr 414
 DB 687 ACCTATGGAAGATGCTTGAGATGCGGTTTCAAGATGTGTCGCCGAAGACCGCAC 746
 QY 415 GluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434
 DB 747 GACCACTTCTGACGCTCTCGAGAGGAGCTGGCGCGAGAACGAGAGAACAGAGGCGCC 806
 QY 435 PheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyTyrPheAspLys 454
 DB 807 TTCTTGGCGGATCTTCAACGAGAGGACTACGACGACATCTCAACGCGCTGGAGCGGAG 866
 QY 455 LeuArgArgThrAlaLysGlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472
 DB 867 CTCAGCGAGCTCGCGCGAGCAGAGTGGCGCTGTTCTATCGCAACAGG 920

RESULT 5

BE231445 713 bp mRNA linear EST 16-OCT-2000
 LOCUS

DEFINITION

S850815 Suaeda salsa ZAP cDNA library Suaeda maritima subsp. salsa
 cDNA similar to c.elegans cDNA yk2b11.3, mRNA sequence.

ACCESSION

BE231445

VERSION

BE231445.1 GI:8996113

KEYWORDS

EST.

SOURCE

Suaeda maritima subsp. salsa

ORGANISM

Suaeda maritima subsp. salsa

REFERENCE

1 (bases 1 to 713)

AUTHORS

Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.

TITLE

Expressed sequence tags from a halophyte Suaeda salsa cDNA library

JOURNAL

Unpublished (2000)

COMMENT

Contact: Hui Zhang

FEATURES

Key Laboratory of Plant Stress Research

source

The Biology Department of Shandong Normal University

1..713

No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC

/organism="Suaeda maritima subsp. salsa"

Tel: (86)531-2960864

/mol_type="mRNA"

Fax: (86)531-2966954

/sub_species="salsa"

Email: zhanghui@sdnu.edu.cn

/db_xref="taxon:126914"

Location/Qualifiers

/clone_lib="Suaeda salsa ZAP cDNA library"

Site 1: EcoRI; Site 2: XhoI; total RNA extraction from

/notes="Organ: aerial part tissue; Vector: lambda zap;

NaCl(400mM) treated Suaeda salsa by RNagent kit(Promega);

/dev_stage="seedling"

mRNA isolation by MESSAGEMAKER kit(GIBCO BRL);

/directional_cDNA_synthesis="Scori XhoI" by cDNA synthesis

kit(STRATAGEN); the ZAP express library by GigapackIII

Gold Cloning kit(STRATAGEN)"

US-10-031-331B-40 (1-473) x BE231445 (1-713)

Alignment Scores:

81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100

Pred. No.:

1 AAAAATGTCAGTATTATGTGCTGATGTGACTTCTCCCACTCTCAGTTTCCACACNT 60

Score:

101 SerLeuAspValIlePheSerAsnTyrLeuMetTyrLeuSerAspGluValGlu 120

Percent Similarity:

61 TCATTGGATGTGATATTCTCAATTTGTTACTCATGTATCTTTCTGATGAGAGGTGGA 120

Best Local Similarity:

121 AsnLeuValGluArgMetLeuLysTyrLeuLysProGlyGlyTyrIlePhePheArgGlu 140

Query Match:

121 GATTGGTTGGAAGAAATGTTGAATGTTGACCCAGGGGGTTATTTTCTTCAGAGNA 180

DB:

141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160

US-10-031-331B-40 (1-473) x BE231445 (1-713)

181 TCTTGTTCATCAATCAATCTGGGATCACAAACGCAAAAGCAATCCACCACTTACCGTGA 240

QY

161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180

DB

241 CTTAGGTTCTACATTAGGCGCTTCAAGAGTGTCAATTTGCMAGATGGATCTCGAATCTCT 300

QY

181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArganLysLysAsn 200

DB

301 TATGAGCTCTCCCTACTTAGCTGCAATGATTATGGAGCTTATGTCAGAAACCAAGAAAC 360

QY

201 GlnAsnGlnIleSerTyrLeuTyrGlnLysValAspSerLysAspAspLysGlyPheGln 220

```

Db      361 CAGACACGATTAGTGTGGTGGCAAAAGATTGATCTTAAGGATGATGAAGGGTCCAG 420
Qy      221 ArgPheLeuAapThrSerGlnTylyeCysAenSerileLeuArgTyrGlu-ArgValph 240
Db      421 CGATTCTCGGATGACTAGCCAGTACAGTGTGTAAGTCAATCTCGCATGATGAAGCGTATT 480
Qy      240 eGlyProGlyTyrValserThrGlyGlyTyrGluThrThrlyeGluPheValserMetle 260
Db      481 TGSCCCTGGTATGTTAGTACCTGCGAGGATATGAACCCACCAAGAGTTTGTCTCAATGCT 540
Qy      260 uAenLeuAapProGlyGlnTylyeValLeuAapValGlyCysGlyileGlyGlyAspPh 280
Db      541 GACCTTGAGCCTCC-CHGAGGTCCTGGAGTGTGGTGTGGGAAATGGTGGAGGTGACTT 599
Qy      280 eTyrMetAlaGluThrPheAapValGluValValGly-PheAapLeuSerValAenMetI 300
Db      600 TTACATGGCGAGACCTTTCATGTTGAGGGTGTGGAAATTTGATCTCTCCGTTAATATGA 659
Qy      300 leSerPheAlaLeu-GluAtrSerileGlyLeuLysCysAlaValGlu 315
Db      660 TTTCCTTTGGCTTTGAGCGTCTATTGGGCTTAAGTGTCTGCTTGAG 707

```

RESULT 6

BM408047
LOCUS
EST582374 potato roots Solanum tuberosum cDNA clone cPRO33G1 5',
end, mRNA sequence.

BM408047
BM408047.1 GI:18259677

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

UNPUBLISHED

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics

Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/

Seq primer: T3.

Location/Qualifiers

1. 767

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="cPRO33G1"

/tissue types="roots"

/dev stage="in vitro grown stem cuttings"

/lab host="SOLR"

/clone.lib="potato roots"

/note=vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; supplier: Cornell University, Tanksley lab;

sequencing; The Institute for Genomic Research. Roots were

isolated from in vitro grown stem cuttings on CM medium.

Roots were isolated two weeks after placing the stem

cuttings from in vitro grown plants on medium."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Length:

Matches:

Conservative:

Best Local Similarity: 82.75%
Query Match: 46.20%
DB: 12
Mismatches: 28
Indels: 0
Gaps: 0

US-10-031-331b-40 (1-473) x BM408047 (1-767)

```

Qy      147 GlyAapHleAaGlyeSerAapProThrHisTyrArgGluProArgPheTyrThrlye 166
Db      1 GAGGACCAAGCAGGAGAGACACCAACCCATTTATCGGAGGCTAGATTTTACACAAAG 60
Qy      167 AlaPheLysGluCysHleLeuGlnAapGlySerGlyAenSerTyrGluLeuSerLeuLeu 186
Db      61 GTGTTTAAAGAAATGTCATATAAATGCTGGTGGTAAATCATTTGAACTTCTCTCATC 120
Qy      187 SerCyLeCysCyellleGlyAlaTyrValArgAenLysAaGlnAaenGlnleSerTyr 206
Db      121 GGTTCGAAGTCATTCGAGCTTAATTTAAACCAACCAAGAAATCAGATTCAGATTTGTGG 180
Qy      207 LeuTTCGlnLysValAapSerLysAapGlyPheGlnArgPheLeuAapThrSer 226
Db      181 CTTGGCAAAAGGTAAATTTCTGAGGATGACAGGGGATTCAGCGTTTCCTGGACAATGT 240
Qy      227 GlnTyrLysCysAenSerileLeuAtrGlyTyrGluArgValPheGlyProGlyTyrValSer 246
Db      241 CAATACAAATGTAGTGGCATACTGCCATATGAACGTGTCTTTGGAGAAAGTTATGTGAGC 300
Qy      247 ThrGlyGlyTyrGluThrThrlyeGluPheValserMetleAapLysProGlyGln 266
Db      301 ACAGGAGGACTTGATACCAACCAAGAAATTCGTTTCTATGTGGATCTTCACCTGGCCAA 360
Qy      267 LysValLeuAapValGlyCysGlyileGlyGlyAapPheTyrMetAlaGluThrPhe 286
Db      361 AAGTCCCTTGATGTGGCTGGAAATAGTGGAGGTGACCTTTTACATGCTGAGAATAT 420
Qy      287 AspValGluValValGlyPheAapLeuSerValAaenMetileSerPheAlaLeuGluArg 306
Db      421 GATGTTTCATGTTGTTGCTATTGACCTTCGATTAAACATGATCGTTGCTCTTGACGT 480
Qy      307 SerileGlyLeuLysCysAlaValGluPheGluValAlaAapCysThrlyleAenTyr 326
Db      481 TCATTTGCTCAATATGCGGTGAAATTTGAGTTGCTGATTTGTACAAAGAAACATAT 540
Qy      327 ProAapAenSerPheAapValileTyrSerArgAapThrileLeuHleGlnAapLys 346
Db      541 CCGTATGGCACAATTTGATGTGATTTACAGTCGAGACACTATCTTCACATCCACAGACAA 600
Qy      347 ProAlaLeuPheAaGSerPheTyrLysTyrLeuLysProGlyGlyLysValleuileSer 366
Db      601 CCCAGTTATTTCAGATCTTCTACAGTGGCTGCTGAAACCCAGAGGCAAGTCTCTATAAGT 660
Qy      367 AspTyrCysLysLysAlaGlyProProSerProGluPheAlaAlaTyrileLysGlnArg 386
Db      661 GATTACTGCAAAATGCCCAATACCAAGCATCGAGAGTTCCTGAGTATATAAAGCAAAAG 720
Qy      387 GlyTyrAapLeuHleAapValLysGluTyrGlyGlnMetLeuLys 401
Db      721 GGTATGATTATACATGATGTTGCAACATATATGCCAGATGCTCAAA 765

```

RESULT 7

CK166076/c

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

CK166076

VERSION

CK166076.1

KEYWORDS

EST.

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poaceae; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 1115)

AUTHORS

Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,

CK166076 1115 bp mRNA linear EST 05-DEC-2003
FGAS050129 Triticum aestivum FGAS: Telt7 Triticum aestivum cDNA,
mRNA sequence.

CK166076

CK166076.1 GI:38998769

EST.

Triticum aestivum (bread wheat)

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poaceae; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 1115)

AUTHORS

Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,

Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A., Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D., Penniket, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: f9ae_scs@usask.ca

TITLE

JOURNAL

COMMENT

This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [65,863].
Plate: Talet707 row: K column: 02.
Location/Qualifiers

FEATURES

source

1. .1115
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="wheat line CI 14106"
/db_xref="taxon:4565"
/lab_host="DH5 alpha"
/clone_lib="Triticum aestivum FGAS: Talet7"
/notes="Organ: Crown; Vector: pGEM-T; SSH (suppression
subtractive hybridization) cDNA library from genotype
CI14106 cold hardened at 2 C for 21 days and 49 days
(equal amount of cDNA pooled together before subtraction,
tester) and subtracted against genotype CI14106
non-hardened (20 C) (driver). Nitro-Pyrole anchored
oligo-dT priming and non-directional cloning."

ORIGIN

Alignment Scores:
Pred. No.: 7.58e-134 Length: 1115
Score: 1158.50 Matches: 225
Percent Similarity: 82.66% Conservative: 42
Best Local Similarity: 69.66% Mismatches: 56
Query Match: 46.06% Indels: 4
DB: 14 Gaps: 0

US-10-031-331B-40 (1-473) x CK166076 (1-1115)

QY 150 LysArgLysSerAnProThrHisTyrArgGluProArgPheTyrThrLysAlaPheLys 169
Db 1110 CAGAGGGGAAATACCGACCCACTACCGGAGCCGAGGTTTACCCACCAAGGTTTAAAG 1051
QY 170 GluCyHisLeuGlnApgLysSerGlyAnSerTyrGluLeuSerLeuLeuSerCysLys 189
Db 1050 GA-TGCCACTCTAGACCA-AGAGGGGATTCCTTGAGCTTCCTCGTAACCTCCAG 993
QY 190 CysLleGlyAlaTyrValArgLysLysLysLysLysLysLysLysLysLysLysLysLys 209
Db 992 TGCAAT-GGAGCTTATGTGAAGCAGGAGCAACCCAGACCCAGATATGTGGCTGGGAG 934
QY 210 LysValaLysSerLysAspLysGlyPheGlnArgPheLysLysLysLysLysLysLys 229
Db 933 AAGGTCAGTCACAGAGACCAAGGCTTCAGAGATCCCTGGACAAATGTCAGTACAA 874
QY 230 CysAsnSerLleLeuArgTyrGluArgValPheGlyProGlyTyrValSerThrGlyGly 249
Db 873 TCCACTGGAAATCTTGCCTATAGCGCTGTGGGAG-GGTTATGTAGACACCGGTGA 815
QY 250 TyrGluThrLysGluPheValSerMetLeuAspLysLysLysLysLysLysLysLys 269
Db 814 TTCGAGACCAAGAGGATTTGTCGACACTCGACTCGAAAGCTGCTCAGAGGTGCTC 755
QY 270 AspValGlyCysGlyLleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 289
Db 754 GATGTAGGCTGTGGTATCGGAGGAGCGGACTTCTACATGGCTGAAACCTATGATGTCCAT 695

QY 290 ValValGlyPheAspLeuSerValaenMetIleSerPheAlaLeuGluArgSerIleGly 309
Db 694 GTCTCGGAGTCACTTCTATCAACATGGTTCGTTCGATTCGATTCGATTCGATTCG 635
QY 310 LeuLysCysAlaValGluPheGluValAlaAspCysThrLysLysLysLysLysLys 329
Db 634 CGCTCGTCTCCGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAG 575
QY 330 SerPheAspValIleTyrSerArgAspThrIleLeuHisLysGlnAspLysProAlaLeu 349
Db 574 ACCTTTGATGTCATCTACAGCCGTCAGACCATCTCCACATTCACAGACAAACCTGCTC 515
QY 350 PheArgSerPheTyrLysTyrLeuLysProGlyLysValLeuLysLysLysLysLys 369
Db 514 TTCAAGAACTTCTCAAGTGGCTCAAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 455
QY 370 LysLysAlaGlyProProSerProGluPheAlaLysLysLysLysLysLysLysLys 389
Db 454 AGGAGCCCTGGGACACCATCAGAGGATTCGCTGTGTACATCAAGCAGAGAGGCTATGAC 395
QY 390 LeuHisAspValLysGlyTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeu 409
Db 394 CTCATGATGTGAAGACCTACCGAAAGATGCTTGAGGATCCGCTTCCATGATGTCGTC 335
QY 410 AlaGluAspArgThrGluGlnPheIleArgValLeuArgLysGluLeuGluThrValGlu 429
Db 334 GCTGAGACCGACCCACCCAGCTTCCTGAGGGTCTCTGAGAGGGAGCTGGGCGAGACGAG 275
QY 430 LysGluLysAspValPheLysSerAspPheSerGluGluAspTyrAsnAspIleValGly 449
Db 274 AAGAACAAAGAGGAGCTTCTCGGAGCTTACCCAGGAGGACTACGACGATCGTCAAC 215
QY 450 GlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGlnArgTyrGlyValPheVal 469
Db 214 GCGCGAGCGGAGTTGAGCGGAGCTTCGCGGAGCAGAGAGTGGGCTGCTTCATC 155
QY 470 AlaLysLys 472
Db 154 GCGACCAAG 146
RESULT 8
CF513471
LOCUS CF513471.1 GI:34545239
DEFINITION Bud - CABUD Vitis vinifera cv. Cabernet sauvignon (Clone 8)
sequence.
ACCESSION CF513471
VERSION CF513471.1
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 869)
AUTHORS Goes da Silva, P., Landolino, A., Lim, H., Baek, J., Jones, K. and Cook, D.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages
Unpublished (2003)
Contact: Douglas Cook, PhD
CARS Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drc@ucdavis.edu
Seq primer: ACGTACCGACATATGCC.
Location/Qualifiers
1. .869
/organism="Vitis vinifera"
/mol_type="mRNA"

PRECEDENT 9

LOCUS	CF208434	781 bp	mRNA	linear	EST 01-AUG-2003
DEFINITION	CAB20003	Iia Fa G05	Cabernet Sauvignon	Flower bloom	- CAB2 Vitis
	vinifera	cDNA clone	CAB20003	Iia Fa G05 5'	mRNA sequence.

Pred. No.:	4,638-128	Length:	869
Score:	111.00	Matches:	202
Percent Similarity:	88.2%	Conservative:	31
Best Local Similarity:	76.5%	Mismatches:	31
Query Match:	44.1%	Indels:	0
DB:	14	Gaps:	0

LOCUS	CF208434	781 bp	mRNA	linear	EST 01-AUG-2003
DEFINITION	CAB20003	Iia Fa G05	Cabernet Sauvignon	Flower bloom	- CAB2 Vitis
	vinifera	cDNA clone	CAB20003	Iia Fa G05 5'	mRNA sequence.

LOCUS	CF208434	781 bp	mRNA	linear	EST 01-AUG-2003
DEFINITION	CAB20003	Iia Fa G05	Cabernet Sauvignon	Flower bloom	- CAB2 Vitis
	vinifera	cDNA clone	CAB20003	Iia Fa G05 5'	mRNA sequence.

Alignment Scores:

LOCUS	CF208434	781 bp	mRNA	linear	EST 01-AUG-2003
DEFINITION	CAB20003	Iia Fa G05	Cabernet Sauvignon	Flower bloom	- CAB2 Vitis
	vinifera	cDNA clone	CAB20003	Iia Fa G05 5'	mRNA sequence.

515	AAAGGCTTTGAGAGATTCTGGACACATGTGCAGGTACAAATCCACTCGGAATCTTGGCGTAC	574
237	Glua-gValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPhe	256
575	GAGCC-GTGTTTGGGGAGGGTTATGTGTAGACACCGGTGGATTTCGAGACCAAAAGGAAATTT	633
257	ValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGly	276
634	GTGGACAAAGCTGGACCTGAAGCCGCCAGAAAGGTGCTTGATGTTCGGTGTGGCATCGGA	699
277	GlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSer	296
694	GGAGGGCACTTCTACATGGCTGAACCTATGATGTCCATGTCTCTGGCATCGATCTTTCC	753
297	ValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPhe	316
754	ATCAACCAATGTTTC-TTCGCGATCGAGCGTGCCATCGTCTGCTTCGTGGTGGTGAAGTTT	812
317	GluValAlaAspCysThrLysIleAsnTyr-ProAspAsnSerPheAspValIleTyrSer	336
813	GAGGGTCTGTGATCGCACCAACCA--GGGATCCCGAGAACACGTTGCATGTCATCTAAAG	870
337	ArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysIrp	356
871	CGTGACCAACAT-CTCCCCATTCAGAACAGCCCTGTCTTTTCAGAAACTCTCTTCAAGTG	929
357	LeuLysProGlyGlyValValLeuIleSerAspTyrCysLysLysAlaGlyProProSer	376
930	GC-AAACTCTGGGGGCAAAAGTCTGGTCGT-GGCTTCTCGAGGACCTGGGGGGCAATCAAG	987
377	ProGluPhe	379
988	GATTCCTTC	996

RESULT 11	BE040284	LOCUS	BE040284	DEFINITION	OB03A05 OB Oriza sativa cDNA 5' similar to s-adenosyl-methionine-sterol-c-methyltransferase, mRNA sequence.	822 bp	mRNA	linear	EST 07-JUN-2000
ACCESSION	BE040284	VERSION	BE040284.1	GI:8335300	EST				
KEYWORDS	ORYZA SATIVA	SOURCE	ORYZA SATIVA	ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
REFERENCE	1 (bases 1 to 822)	AUTHORS	Bohner, H.J., Borchert, C., Brazille, S., Brooks, J., Baton, M., Ferreira, H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C., Scaras, G., Wheeler, M. and Zepeda, G.R.						
TITLE	Functional Genomics of Plant Stress Tolerance	JOURNAL	Unpublished (2000)						
COMMENT	Contact: Michalowski, C.B.		University of Arizona						
	Bio Sciences West room 513, Tucson, AZ 85721, USA								
	Tel: 520-621-7982								
	Fax: 520-621-1697								
	Email: cbm@u.arizona.edu								
	An open reading frame exists.								
	Insert Length: 1								
	Std Error: 0.00.								

FEATURES
SOURCE

ORIGIN

RESULT 12	ACCESSION
BQ965516/c	VERSION
LOCUS	
DEFINITION	

Alignment Scores:
Pred. No.: 1,9e-122 Length: 822
Score: 1066.00 Matches: 205
Percent Similarity: 85.40% Conservative: 29
Best Local Similarity: 74.82% Mismatches: 40
Query Match: 42.39% Indels: 3
DB: 10 Gaps: 0

US-10-031-331B-40 (1-473) x BE040284 (1-822)

QY	159	ArgGluProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGly	178
DB	1	CGGGAGCCAGAGTTTGAGACTAAGTGTTTAAAGAGTGTCAGAGCTCTTGATCAAGATGGG	60
QY	179	AnsSerTyrGluLeuSerLeuLeuSerCysLysCysIleGlyValatyrValargAsnLys	198
DB	61	AATTCCTTTGAATCTCTGTACTTACTTGCNAGTGTTGGAGCTTAAGTGAAGACCAAG	120
QY	199	LysAsnGlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspGlyGly	218
DB	121	AAANATCAAAACCAGATATGTTGGCTATGSCAAAGAGTTGATTCACAGAGATCGGGG	180
QY	219	PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg	238
DB	181	TTTCAAGATTTTGGCAATGTGCAGTACAAAGCCAGTGGATATATAGCTATGACGC	240
QY	239	ValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSer	258
DB	241	ATCTTTGGAGAGGCTTTGTGGACCACTGGTGGAAATGAAACTACAAAGAAATTTGTGGAC	300
QY	259	MetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyGly	278
DB	301	AGGCTGGATCTCAAACTGGCCAGAACGTTCTTGATGTGGATGTGGAAATGGGGGGCGGT	360
QY	279	AspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsn	298
DB	361	GATTTTATATGGCTGACAGATATGATGTTTCATGTTGTTGGTATTGATCTTTCGATAAAC	420
QY	299	MetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluVal	318
DB	421	ATGGTTTCTTTTGACATTGACGGTGCTATTGGGGCGTAAAGTCTCAGTGTGTTGAGTCT	480
QY	319	AlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAsp	338
DB	481	GCTGATGTCACCAAAAGACATATCCACAGACACACGTTTGACGTCACTATCACAGTCGTAT	540
QY	339	ThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLys	358
DB	541	ACTATCTTCACATACAGATATAACCTCTACTATTTAAAGTTTCTTCAGTGGGCTCAA	600
QY	359	ProGlyGlyLysValLeuIleSerAspTyrCysGlyLysValaGlyProProSerProGlu	378
DB	601	CTGTGGGGTAGGTGCTCTAATAGTAGATATCTGCAAGTGCCCTGGGAAACCTTCAAGAGAG	660
QY	379	PheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGln	398
DB	661	TTCCGAGCTTACATTAAAGCAAAAGGGTATGACCTTAACCCAGCTCAGGCGCTACGACAG	720
QY	399	MetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIle	418
DB	721	ATGCTTGA-AAATGCTGGTTTCCATGTATGATTCATGGTGAAGACCGCAACCTCAGTTCCT	778
QY	419	ArgValLeuArgLysGluLeuGluThrValGluLysGluLys	432
DB	779	CGAGTTCTAGAGAGGAGCCTGGT-AAAGTTTGAAGACCAAAA	819

BQ965516 733 bp mRNA linear EST 21-AUG-2002
QH822C02.yf.ab1 QH ABCDI sunflower RH801 Helianthus annuus cDNA
clone QH822C02, mRNA sequence.

[illegible]


```

Db      423 GTTTCATTGCAATCGAAGCGTCCATTGGACGCAAGTGTTCGGTTGAGTTGAAGTTGCT 482
Qy      320 AspCythrIlyleAsnTyrProAspSerPheAspValIleTyrSerArgAspThr 339
Db      483 GATTGACCAACAAAGACCTAGCCACCAATATCTTGTGATGTGATCTACAGCGGTGACACC 542
Qy      340 IleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrIlyTrpLeuLysPro 359
Db      543 ATTCTTCACATACATATAACCTGCTNTGTTTCAGAGTTTCTTCAAGTGGCTGAAACCT 602
Qy      360 GlyIlyLysValLeuIleSerAspTyrCysIlyLysAlaGlyProProSerProGluPhe 379
Db      603 GNGGCAAGTCTCTCATCTAGTATCTAGCAATCTCTGGAAACCATCAGAGAATTT 662
Qy      380 AlaAlaTyrIlyLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMet 399
Db      663 GCTGCTTACATTAAGCGAGAGGCTATGACCTCCAGATGTGAGACTTACCGAGAGATG 722
Qy      400 LeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArg 419
Db      723 CTTGAGGATCTGGGGTTTCATCATGCAATGCTGTAAGACGC-ACGACCAAGTCTCTGCGT 781
Qy      420 ValLeuArgLysGluLeu 425
Db      782 GTTCTTCAAGAGGCTTG 799

RESULT 15
CF482195
LOCUS   721 bp mRNA linear EST 08-SEP-2003
DEFINITION POL1_5_B02_g1_A002 Pollen Sorghum bicolor cDNA clone
ACCESSION POL1_5_B02_A002 5', mRNA sequence.
VERSION   CF482195
KEYWORDS  EST.
SOURCE    Sorghum bicolor (sorghum)
ORGANISM  Sorghum bicolor

REFERENCE
AUTHORS  Cordonnier-Pratt, M.-M., Suzuki, Y., Sugano, S., Klein, R.R., Liang, C.,
Sun, F., Sullivan, R., Eastman, A., Cannon, R., Kern, B., Morgan, J.,
Lucas, A., Al-Sheikh, A., Jones, V., Adibi, N., Owen, A., Gao, J. and
Pratt, L.H.
EST database from Sorghum: pollen
Unpublished (2003)
Other ESTs: POL1_5_B02_b1_A002
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCGTCTCTAAAGCTGGC).
FEATURES             Location/Qualifiers
     source            1..721
     organism="Sorghum bicolor"
     mol_type="mRNA"
     cultivar="BTx623"
     db_xref="taxon:4558"
     clone="POT1-9 B02_A002"
     lab_host="DH10B-T1 phage-resistant E. coli"
     clone_lib="Pollen"
     note="Organ: Pollen; Vector: pME18S-FL3; Site_1: XhoI;

```

Site 2: XhoI; The library was prepared from polyA+ RNA from pollen at the late vacuolated-vacuolated stage of development. Pollen was harvested from greenhouse-grown panicles of sorghum line BTx623. Panicles were removed from the flag leaf prior to emergence, when no detectable amylose is present in pollen of male-fertile lines. This stage represents pollen collected from anthers about 8-14 days prior to anthesis. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5'-prime DraIII site is CACTGTGTG, 3'-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

```

Alignment Scores:
Score: 5.02e-119 Length: 721
Percent Similarity: 1038.00 Matches: 190
Best Local Similarity: 89.92% Conservatives: 24
Query Match: 79.83% Mismatches: 24
DB: 41.27% Indels: 0
Gaps: 14

US-10-031-331B-40 (1-473) x CF482195 (1-721)

Qy      235 ArgTyrGluArgValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLys 254
Db      8 CGTTATGAGCGGTGCTTTGTGGAAGGTTATGTGAGCACTGTGTGGAAATCGAGACTACAAAG 67
Qy      255 GluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGly 274
Db      68 GAATTTGTGGCATGCTGGATCTTAAACCTGGCAAGAAAGTACTTGATGTGGATGGGA 127
Qy      275 IleGlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAsp 294
Db      128 ATTGAGCGCGGCACTTTACATGCTGGAACATATGATGCTCATGCTGTGGATTGAT 187
Qy      295 LeuSerValSerMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaVal 314
Db      188 CTTTGGGTTAAACATGTTTCATTTGCAATGAACGTCGTCATTTGAGAGCTTCTTC 247
Qy      315 GluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIle 334
Db      248 GAATTTGAGTTGCTGATTGTCACCAACAAAGGATTATCCCAAGAGATAGTTTTCAGCTCATC 307
Qy      335 TyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyr 354
Db      308 TACAGCGGTGACACTATCTTTCATACAGACAGACGCTGCTCTGTTGAGAGCTTCTTC 367
Qy      355 LysTrpLeuLysProGlyGlyLysValLeuIleSerAspTyrCysLysLysAlaGlyPro 374
Db      368 AAATGGCTTAAACCCCGTGGCAAGTCTTAATCAGTCACTACTGTAGAAATCTCTGAAAAA 427
Qy      375 ProSerProGluPheAlaAlaTyrIleLysValGlnArgGlyTyrAspLeuHisAspValLys 394
Db      428 CCATCAGAGAAATTTGCTGATACATATAGCAGAGAGGTTATGACCTCCACGATGTGAAG 487
Qy      395 GluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThr 414
Db      488 GCTTATGACAGATGCTCAAGGATGCTGTTTCATGATGTCTATGATGTCTGAGAGATGCGACT 547
Qy      415 GluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434
Db      548 GAGCAGTTCTGTAATGTTCTACCGAGGAGAGTAGTGGAAGTTGAAAGACAAAGAGGCT 607
Qy      435 PheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTrpAsnAspLys 454
Db      608 TTCTTGGGAGACTTTCACCCAGGAGGACTACGACGACATTTGTAACGCTGGAACGCGAAG 667
Qy      455 LeuArgArgThrAlaLysGlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
Db      668 CTGAACCGGAGCTCTGCGCGGTGAGCAGAGGTGGGGCTCTGTTTCATTGGACCAAG 721

```

Tue Aug 3 10:37:38 2004

us-10-031-331b-40.rst

Page 14

Search completed: August 2, 2004, 03:21:23
Job time : 3593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:15:57 ; Search time 54 Seconds
(without alignments)
2474.907 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 473

Sequence: 1 HTVDLTIEAWMLDSQASLDL.....KLRTAKGEQRWGLFVAKKK 473

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	473	100.0	473	4	AAB80627
2	45	9.5	289	4	AAB09761
3	45	9.5	494	4	AAB09760
4	23	4.9	191	7	ABM74064
5	21	4.4	168	4	ABM74064
6	21	4.4	491	4	ABM74064
7	16	3.4	287	7	ABM74064
8	16	3.4	346	5	AAE13618
9	14	3.0	293	3	AAE13618
10	14	3.0	293	3	AAE13618
11	14	3.0	293	3	AAE13618
12	14	3.0	293	3	AAE13618
13	14	3.0	314	5	ABM74064
14	14	3.0	336	3	AAE13618
15	14	3.0	336	3	AAE13618
16	14	3.0	336	3	AAE13618
17	14	3.0	339	3	AAE13618
18	14	3.0	358	3	AAE13618
19	14	3.0	361	3	AAE13618
20	14	3.0	365	3	AAE13618
21	14	3.0	366	3	AAE13618
22	11	2.3	137	3	AAE13618
23	11	2.3	140	3	AAE13618
24	11	2.3	152	3	AAE13618
25	11	2.3	202	6	ABO01277

26	11	2.3	344	2	AAW70838
27	11	2.3	344	3	AAW70838
28	10	2.1	301	4	AAW70838
29	10	2.1	309	6	AAW70838
30	10	2.1	318	3	AAW70838
31	10	2.1	318	4	AAW70838
32	10	2.1	318	4	AAW70838
33	10	2.1	318	5	AAW70838
34	10	2.1	318	6	AAW70838
35	9	1.9	280	6	AAW70838
36	9	1.9	280	6	AAW70838
37	9	1.9	280	7	AAW70838
38	9	1.9	280	7	AAW70838
39	9	1.9	311	6	AAW70838
40	9	1.9	330	6	AAW70838
41	9	1.9	333	4	AAW70838
42	9	1.9	621	4	AAW70838
43	9	1.9	653	4	AAW70838
44	8	1.7	43	4	AAW70838
45	8	1.7	43	6	AAW70838

ALIGNMENTS

RESULT 1

AAB80627
ID AAB80627 standard; protein; 473 AA.

AC AAB80627;

DT 06-AUG-2003 (revised)

DT 02-MAY-2001 (first entry)

DE Environmental stress tolerant protein SEQ ID 40.

XX Environmental stress resistance; salt; heat; desert; transgenic plant.

OS Suaeda japonica.

PN WO200106006-A1.

PD 25-JAN-2001.

PF 19-JUL-2000; 2000WO-JP004862.

PR 19-JUL-1999; 99JP-00235910.

PR 24-MAR-2000; 2000JP-00085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Yamada A, Ozeki Y, Saito T;

XX WPI, 2001-147355/15.

XX N-PSDB; AAF74206.

Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

Claim 64; Page 125-127; 167pp; Japanese.

Poly-nucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance factors. The DNA encoding proteins conferring environmental stress resistance, can be used in the production of plants resistant to environmental stress, which can be cultivated in unfavourable

CC environments such as deserts, salt damaged ground, cold regions and the
 CC oceans. They can be used for increasing the area of land covered by green
 CC plants, and desert greening and afforestation, in order to counter the
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR
 CC primers AAF74219 and AAF74220 are used in an example illustrating the
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)
 XX

XX SQ Sequence 473 AA;

Query Match 100.0%; Score 473; DB 4; Length 473;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HTVLTTEAMMLDSQASDLDERPEILSMLEKCKLLEAGIGRTGELAKAGQV 60
 DB 1 HTVLTTEAMMLDSQASDLDERPEILSMLEKCKLLEAGIGRTGELAKAGQV 60
 QY 61 IALDPTESAIKONEVINGHYKVNKPCADVTSPTLSPPHSLDVI FSNWLLMYLSDBRVE 120
 DB 61 IALDPTESAIKONEVINGHYKVNKPCADVTSPTLSPPHSLDVI FSNWLLMYLSDBRVE 120
 QY 121 NLVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFYTKAFKECHLQDGGNS 180
 DB 121 NLVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFYTKAFKECHLQDGGNS 180
 QY 181 YELSLSCCKICAVRNKQNOISLWQKVDSDKQFORPLDTSQYKCNLSLYRVP 240
 DB 181 YELSLSCCKICAVRNKQNOISLWQKVDSDKQFORPLDTSQYKCNLSLYRVP 240
 QY 241 GPGVSTGCGYTTKEFVSMLEKPGQKVLVDGCGIGGDFYMAETFDVEVVGFDLSVNM 300
 DB 241 GPGVSTGCGYTTKEFVSMLEKPGQKVLVDGCGIGGDFYMAETFDVEVVGFDLSVNM 300
 QY 301 SFALERSI GLKCAVEFEVADCTKINYPDNSFDVYSRDTILHI QKPALEFSFYKWLKPG 360
 DB 301 SFALERSI GLKCAVEFEVADCTKINYPDNSFDVYSRDTILHI QKPALEFSFYKWLKPG 360
 QY 361 KGVLSIDYCKKAGPPSPPEPAAIKQGYDLHDVHYGQMLKQAGFVDVLADRTQPTIRV 420
 DB 361 KGVLSIDYCKKAGPPSPPEPAAIKQGYDLHDVHYGQMLKQAGFVDVLADRTQPTIRV 420
 QY 421 LRKELETVEKQVFI SDFSEEDYNDIVGWNDKLARTAGQORWGLFVAKKK 473
 DB 421 LRKELETVEKQVFI SDFSEEDYNDIVGWNDKLARTAGQORWGLFVAKKK 473

RESULT 2

AAE09761
 ID AAE09761 standard; protein; 289 AA.

XX AC AAE09761;

XX DT 29-NOV-2001 (first entry)

XX DE Spinach PEAMT truncated protein.

XX KW Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
 KW glycine betaine; choline-O-sulphate; lipid content alteration;
 KW osmotic stress tolerance; nutritional value; transgenic plant;
 KW cryoprotectant.

XX OS Spinacia oleracea.

XX PN WO200168870-A2.

XX PD 20-SEP-2001.

XX PF 15-MAR-2001; 2001WO-US008352...

XX PR 15-MAR-2000; 2000US-00525885.

XX XX

PA (UYFL) UNIV FLORIDA.
 PA (UYCA-) UNIV CARNEGIE MELLON.

XX PI Hanson AD, Nuccio ML, Henry SA;

XX DR WPI; 2001-565796/63.

XX DR N-PSDB; AAD16798.

XX PT New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
 PT polypeptides, useful for modulating the levels of cellular intermediates
 PT such as phosphodimethylethanolamine and for altering the lipid content in
 PT plants cells.

XX PS Claim 1; Page 109; 158pp; English.

XX CC The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-
 CC methyltransferase (PEAMT) truncated protein from spinach. The PEAMT
 CC sequences are useful for modulating the levels of cellular intermediates
 CC such as phosphodimethylethanolamine, phosphomono-methylethanolamine,
 CC choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or
 CC glycine betaine. They are useful for altering the lipid content in plant
 CC cells. The polynucleotides are also useful for improving the osmotic
 CC stress tolerance of a plant and increasing the cryoprotectant properties
 CC of a plant. The present invention also relates to methods and
 CC compositions comprising PEAMT used for generating transgenic plants with
 CC increased nutritional value

XX SQ Sequence 289 AA;

Query Match 9.5%; Score 45; DB 4; Length 289;
 Best Local Similarity 100.0%; Pred. No. 8.1e-40;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 LVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFYTK 166
 DB 143 LVERMLKWLKPGGYIFFRESCFHQSDHKKSNPTHYREPRFYTK 187

RESULT 3

AAE09760

ID AAE09760 standard; protein; 494 AA.

XX AC AAE09760;

XX DT 29-NOV-2001 (first entry)

XX DE Spinach PEAMT protein.

XX KW Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
 KW glycine betaine; choline-O-sulphate; lipid content alteration;
 KW osmotic stress tolerance; nutritional value; transgenic plant;
 KW cryoprotectant.

XX OS Spinacia oleracea.

XX PN Key Location/Qualifiers

XX FT MISC-difference 462

XX FT Note= "This residue is given as Lys in the sequence
 FT shown as SEQ ID NO: 2 in figure 3 of the specification"

XX PN WO200168870-A2.

XX PD 20-SEP-2001.

XX PF 15-MAR-2001; 2001WO-US008352.

XX PR 15-MAR-2000; 2000US-00525885.

XX PA (UYFL) UNIV FLORIDA.

XX PA (UYCA-) UNIV CARNEGIE MELLON.

XX XX

PI Hanson AD, Nuccio ML, Henry SA;

XX WPI; 2001-565796/63.
DR N-PSDB; AAD16797.

XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
PT polypeptides, useful for modulating the levels of cellular intermediates
PT such as phosphodimethylmethanolamine and for altering the lipid content in
PT plants cells.

XX Claim 1; Page 109; 158pp; English.

XX The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-
CC methyltransferase (PEMT) protein from spinach. The PEMT sequences are
CC useful for modulating the levels of cellular intermediates such as
CC phosphodimethylmethanolamine, phosphomono-methylmethanolamine, choline,
CC phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine
CC betaine. They are useful for altering the lipid content in plant cells.
CC The polynucleotides are also useful for improving the osmotic stress
CC tolerance of a plant and increasing the cryoprotectant properties of a
CC plant. The present invention also relates to methods and compositions
CC comprising PEMT used for generating transgenic plants with increased
CC nutritional value

XX SQ Sequence 494 AA;

Query Match 9.5%; Score 45; DB 4; Length 494;
Best Local Similarity 100.0%; Pred. No. 1.4e-39;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 LVERMLKPGGVIFFRESCFQSGDHKKSNTHTYRPRFTK 166
DB 143 LVERMLKPGGVIFFRESCFQSGDHKKSNTHTYRPRFTK 187

RESULT 4

ABM74064
ID ABM74064 standard; protein; 191 AA.

AC ABM74064;

XX 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP sequence #474.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.

XX Hordeum vulgare.

XX WO2003057877-A1.

XX 17-JUL-2003.

XX 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

XX 20-DEC-2001; 2001JP-00387131.

XX 20-DEC-2001; 2001JP-00403299.

XX 20-DEC-2001; 2001JP-00403300.

XX 27-SEP-2002; 2002JP-00327515.

XX (UTNI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA

XX sequences containing them for analysis and identification of barley

XX varieties and production of barley transformants with desired

XX characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX

CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences

XX SQ Sequence 191 AA;

Query Match 4.3%; Score 23; DB 7; Length 191;
Best Local Similarity 100.0%; Pred. No. 5.4e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 FDVIYSRDTLHLDKPKALFRSF 353
DB 50 FDVIYSRDTLHLDKPKALFRSF 72

RESULT 5

AAB99873

ID AAB99873 standard; protein; 168 AA.

XX AAB99873;

XX 21-SEP-2001 (first entry)

XX Physcomitrella patens 78_pppprotl_092_e12rev protein.

XX Tocopherol and carotenoid metabolism related protein; TCWRP; synthesis;
XX Physcomitrella patens; moss; algae; microorganism; fungus; plant;
XX Identification; genome mapping; modulation; evolutionary study;
XX cellular production; fine chemical.

XX Physcomitrella patens.

XX WO200144276-A2.

XX 21-JUN-2001.

XX 14-DEC-2000; 2000WO-EP012698.

XX 16-DEC-1999; 99US-0171121P.

XX (BADI) BASF PLANT SCI GMBH.

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Blaschoff F;
XX Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;

XX WPI; 2001-398121/42.

XX N-PSDB; AAH44246.

XX Tocopherol and carotenoid metabolism related protein (TCWRP), used to
XX produce fine chemicals, is isolated from mosses, algae, microorganisms,
XX fungi, plants, or their fragments.

XX Claim 28; Page 118-119; 123pp; English.

XX The present invention describes isolated tocopherol and carotenoid
XX metabolism related proteins (TCWRP) (I) from mosses or algae,
XX microorganisms or fungi, plants, or its fragments. (I) can be used as
XX enzymes in the production of fine chemicals or in the metabolism of
XX tocopherols and carotenoids. (I) also assist in transmembrane transport.
XX The fine chemicals that can be produced include lipids, fatty acids,
XX vitamins, cofactors, enzymes, amino acids, and nucleotide bases.
XX Nucleotide sequences, proteins, vectors and host cells from the present
XX invention can be used: (a) to identify mosses related to Physcomitrella
XX patens; (b) in mapping genomes of mosses related to Physcomitrella patens

CC : (c) in the modulation of TCMP activity; (d) in evolutionary studies;
 CC (e) in the determination of functional TCMP regions; (f) and in the
 CC cellular production of fine chemicals. AAH44222 to AAH44262 encode the
 CC Physcomitrella patens TCMP proteins given in AAB99849 to AAB99889.
 CC AAH44212 to AAH44221 represent nucleotide sequence used in the
 CC exemplification of the present invention

XX SQ Sequence 168 AA;
 Query Match 4.4%; Score 21; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 7.3e-14;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 SFQVYSKDTILHIOKQPALE 350
 DB 24 SFQVYSKDTILHIOKQPALE 44

RESULT 6
 AAB99889
 ID AAB99889 standard; protein; 491 AA.
 AC AAB99889;
 XX 21-SEP-2001 (first entry)
 DT
 DE Physcomitrella patens 78_ppprot1_092_e12-260rev protein.
 XX
 KW Tocopherol and carotenoid metabolism related protein; TCMP; synthesis;
 KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;
 KW identification; genome mapping; modulation; evolutionary study;
 KW cellular production; fine chemical.
 XX Physcomitrella patens.
 OS WO200144276-A2.
 XX 21-JUN-2001.
 PD
 PF 14-DEC-2000; 2000WO-EP012698.
 XX 16-DEC-1999; 99US-0171121P.
 PR (BADI) BASF PLANT SCI GMBH.
 PA Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff P;
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;
 XX WPI; 2001-398121/42.
 DR N-PSDB; AAH44262.
 XX
 CC Tocopherol and carotenoid metabolism related protein (TCMP), used to
 CC produce fine chemicals, is isolated from mosses, algae, microorganisms,
 CC fungi, plants, or their fragments.
 CC Claim 28; Page 122-123; 123pp; English.
 XX
 CC The present invention describes isolated tocopherol and carotenoid
 CC metabolism related proteins (TCMP) (I) from mosses or algae,
 CC microorganisms or fungi, plants, or its fragments. (I) can be used as
 CC enzymes in the production of fine chemicals or in the metabolism of
 CC tocopherols and carotenoids. (I) also assist in transmembrane transport.
 CC The fine chemicals that can be produced include lipids, fatty acids,
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.
 CC Nucleotide sequences, proteins, vectors and host cells from the present
 CC invention can be used; (a) to identify mosses related to Physcomitrella
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella patens
 CC ; (c) in the modulation of TCMP activity; (d) in evolutionary studies;
 CC (e) in the determination of functional TCMP regions; (f) and in the
 CC cellular production of fine chemicals. AAH44222 to AAH44262 encode the
 CC Physcomitrella patens TCMP proteins given in AAB99849 to AAB99889.
 CC AAH44212 to AAH44221 represent nucleotide sequence used in the
 CC exemplification of the present invention

XX SQ Sequence 491 AA;
 Query Match 4.4%; Score 21; DB 4; Length 491;
 Best Local Similarity 100.0%; Pred. No. 2.1e-13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 PGGYIPFRSCFHSGDHKKK 152
 DB 148 PGGYIPFRSCFHSGDHKKK 168

RESULT 7
 ABM74462
 ID ABM74462 standard; protein; 287 AA.
 XX
 AC ABM74462;
 XX 17-OCT-2003 (first entry)
 DT
 DE DNA clone originating in barley containing SNP sequence #872.
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
 KW Hordeum vulgare.
 OS WO2003057877-A1.
 XX
 PD 17-JUL-2003.
 PF 16-DEC-2002; 2002MO-IB005403.
 XX 20-DEC-2001; 2001JP-00387059.
 PR 20-DEC-2001; 2001JP-00387131.
 PR 20-DEC-2001; 2001JP-00403299.
 PR 20-DEC-2001; 2001JP-00403300.
 PR 27-SEP-2002; 2002JP-00327315.
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 XX Sato K, Takeda K, Kohara Y;
 XX WPI; 2003-587127/55.
 XX
 CC Single nucleotide polymorphism sites in barley varieties and DNA
 CC sequences containing them for analysis and identification of barley
 CC varieties and production of barley transformants with desired
 CC characteristics.
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 XX
 CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences

XX SQ Sequence 287 AA;
 Query Match 3.4%; Score 16; DB 7; Length 287;
 Best Local Similarity 100.0%; Pred. No. 3.5e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 IFSNLLMYLSDEEVE 120
 DB 141 IFSNLLMYLSDEEVE 156

```

RESULT 8
AAE13618
ID AAE13618 standard; protein; 346 AA.
XX AC AAE13618;
XX DT 07-MAR-2002 (first entry)
XX DE Nicotiana tabacum smt1 (Ntsmt1) full length protein.
XX KW Sterol methyl transferase 1; SMT1 gene; cholesterol; oilseed; tobacco.
XX OS Nicotiana tabacum.
XX PN WO200179513-A2.
XX PD 25-OCT-2001.
XX PP 04-APR-2001; 2001WO-EP003808.
XX PR 14-APR-2000; 2000EP-00303193.
XX PA (UNIL ) UNILEVER NV.
XX PA (UNIL ) UNILEVER PLC.
XX PA (UNIL ) HINDUSTAN LEVER LTD.
XX PI Harker M, Hellyer SA, Holmber N, Safford D;
XX DR WPI: 2002-034363/04.
XX DR N-PSDB; AAD23672.
XX PT Using a gene expressing sterol methyl transferase 1 to increase the level
XX PT of sterols in plant seeds or decrease cholesterol level in plant tissue
XX PT is useful to produce commercially desired plants.
XX PS Example 1; Fig 1B; 50pp; English.
XX CC The invention relates to the use of a gene expressing sterol methyl
XX CC transferase 1 (SMT1) to increase the level of sterols in plant seeds or
XX CC decrease cholesterol level in plant tissue. SMT1 genes of the invention
XX CC are used to produce commercially desired plants and oilseeds where the
XX CC plant tissue contains reduced cholesterol and the oilseeds contain
XX CC sterols that may be extracted along with the oils. The present sequence
XX CC is Nicotiana tabacum smt1 (Ntsmt1) full length protein
XX SQ Sequence 346 AA;
Query Match 3.4%; Score 16; DB 5; Length 346;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 262 LKPGQKVLGVGCGIGG 277
Db 98 LKPGQKVLGVGCGIGG 113
RESULT 9
AAG45955
ID AAG45955 standard; protein; 293 AA.
XX AC AAG45955;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 57758.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; Genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.

```

```

PN XX EP1033405-A2.
PD XX
XX XX 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 18-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
XX PR 25-MAY-1999; 99US-0136021P.
XX PR 27-MAY-1999; 99US-0136392P.
XX PR 28-MAY-1999; 99US-0136782P.
XX PR 01-JUN-1999; 99US-0137222P.
XX PR 03-JUN-1999; 99US-0137528P.
XX PR 04-JUN-1999; 99US-0137502P.
XX PR 07-JUN-1999; 99US-0137744P.
XX PR 08-JUN-1999; 99US-0138034P.
XX PR 10-JUN-1999; 99US-0138540P.
XX PR 10-JUN-1999; 99US-0138847P.
XX PR 14-JUN-1999; 99US-0139119P.
XX PR 16-JUN-1999; 99US-0139452P.
XX PR 16-JUN-1999; 99US-0139453P.
XX PR 17-JUN-1999; 99US-0139452P.
XX PR 18-JUN-1999; 99US-0139454P.
XX PR 18-JUN-1999; 99US-0139455P.
XX PR 18-JUN-1999; 99US-0139456P.
XX PR 18-JUN-1999; 99US-0139457P.
XX PR 18-JUN-1999; 99US-0139458P.
XX PR 18-JUN-1999; 99US-0139459P.
XX PR 18-JUN-1999; 99US-0139460P.
XX PR 18-JUN-1999; 99US-0139461P.
XX PR 18-JUN-1999; 99US-0139462P.
XX PR 18-JUN-1999; 99US-0139463P.
XX PR 18-JUN-1999; 99US-0139750P.
XX PR 18-JUN-1999; 99US-0139763P.
XX PR 21-JUN-1999; 99US-0139817P.
XX PR 22-JUN-1999; 99US-0139899P.
XX PR 23-JUN-1999; 99US-0140353P.
XX PR 23-JUN-1999; 99US-0140354P.
XX PR 24-JUN-1999; 99US-0140695P.
XX PR 28-JUN-1999; 99US-0140823P.
XX PR 28-JUN-1999; 99US-0140891P.
XX PR 30-JUN-1999; 99US-0141287P.
XX PR 01-JUL-1999; 99US-0141842P.

```

PR 01-JUL-1999; 99US-01421154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142330P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144353P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144684P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145129P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145221P.
 PR 23-JUL-1999; 99US-0145222P.
 PR 26-JUL-1999; 99US-0145227P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 28-JUL-1999; 99US-0145915P.
 PR 28-AUG-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148563P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151063P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159310P.
 PR 14-OCT-1999; 99US-0159311P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161933P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 293;

Best Local Similarity 100.0%; Pred. No. 5.5e-06;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 PGQKVLVGGGIGG 277

DB 51 PGQKVLVGGGIGG 64

RESULT 10

AAG21652

ID AAG21652 standard; protein; 293 AA.

XX AC AAG21652;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 24284.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KM hybridisation assay; Genetic mapping; gene expression control; promoter;

XX KM termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135112P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135639P.
PR 25-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136382P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 14-JUN-1999; 99US-0138477P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140693P.
PR 28-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 02-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-014320P.
PR 13-JUL-1999; 99US-0143977P.
PR 99US-01439542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-01455913P.
PR 27-JUL-1999; 99US-01455918P.
PR 28-JUL-1999; 99US-01455919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151086P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 26-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132853P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135639P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140981P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.

```
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 296;
Best Local Similarity 100.0%; Pred. No. 5.5e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

QY 264 PGQKVLVDGCGIGG 277
DB 51 PGQKVLVDGCGIGG 64

RESULT 13
ABB93454
ID ABB93454 standard; protein; 314 AA.
AC ABB93454;
XX
XX 31-MAY-2002 (first entry)
XX
XX Herbicidally active polypeptide SEQ ID NO 2665.
XX
XX Herbicidal; plant; agriculture; herbicide.
XX
XX Arabidopsis thaliana.
XX
XX WO200210210-A2.
XX
XX 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP009892.
XX
XX 28-AUG-2001; 2001WO-EP009892.
XX
XX (FARB ) BAYER AG.
XX
XX Tietjen K, Weidler M;
XX
XX WPI; 2002-369010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX
XX Claim 5; SEQ ID NO 2665; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides

PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 5; Length 314;
Best Local Similarity 100.0%; Pred. No. 5.8e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

QY 264 PGQKVLVDGCGIGG 277
DB 72 PGQKVLVDGCGIGG 85

RESULT 14
RAG21651
ID RAG21651 standard; protein; 336 AA.
XX
XX AC RAG21651;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 24283.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX
XX 05-MAR-1999; 99US-0123180P.
XX
XX 23-MAR-1999; 99US-0123548P.
XX
XX 23-MAR-1999; 99US-0125788P.
XX
XX 29-MAR-1999; 99US-0126264P.
XX
XX 01-APR-1999; 99US-0126785P.
XX
XX 08-APR-1999; 99US-0127462P.
XX
XX 16-APR-1999; 99US-0128234P.
XX
XX 16-APR-1999; 99US-0128714P.
XX
XX 19-APR-1999; 99US-0129845P.
XX
XX 21-APR-1999; 99US-0130077P.
XX
XX 23-APR-1999; 99US-0130449P.
XX
XX 23-APR-1999; 99US-0130510P.
XX
XX 28-APR-1999; 99US-0130891P.
XX
XX 30-APR-1999; 99US-0131449P.
XX
XX 04-MAY-1999; 99US-0132407P.
XX
XX 05-MAY-1999; 99US-0132408P.
XX
XX 06-MAY-1999; 99US-0132485P.
XX
XX 06-MAY-1999; 99US-0132486P.
XX
XX 07-MAY-1999; 99US-0132487P.
XX
XX 11-MAY-1999; 99US-0132863P.
XX
XX 14-MAY-1999; 99US-0134256P.
XX
XX 14-MAY-1999; 99US-0134218P.
XX
XX 14-MAY-1999; 99US-0134219P.
XX
XX 14-MAY-1999; 99US-0134221P.
XX
XX 18-MAY-1999; 99US-0134370P.
XX
XX 19-MAY-1999; 99US-0134768P.
XX
XX 20-MAY-1999; 99US-0134941P.
XX
XX 21-MAY-1999; 99US-0135124P.
XX
XX 24-MAY-1999; 99US-0135333P.
XX
XX 25-MAY-1999; 99US-0135629P.
XX
XX 27-MAY-1999; 99US-0136021P.
XX
XX 28-MAY-1999; 99US-0136392P.
XX
XX 01-JUN-1999; 99US-0136782P.
XX
XX 03-JUN-1999; 99US-0137222P.
XX
XX 04-JUN-1999; 99US-0137528P.
XX
XX 07-JUN-1999; 99US-0137502P.
XX
XX 07-JUN-1999; 99US-0137724P.
```

PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140823P.
PR 26-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144634P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144881P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151030P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160816P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match3.0%; Score:14;...DB 3; Length 336;
Best Local Similarity 100.0%; Pred. NO. 6.2e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 264 PCKKLVGGGIGG 277
DB 94 PCKKLVGGGIGG 107

RESULT 15
AAG45951
ID AAG45951 standard; protein; 336 AA.
XX AAG45951;
AC AAG45951;
XX AAG45951;
DT 18-OCT-2000 (first entry)
XX AAG45951;
DE Arabidopsis thaliana protein fragment SEQ ID NO: 57753.
XX Arabidopsis thaliana protein fragment SEQ ID NO: 57753.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX EP1033405-A2.
PD 06-SEP-2000.
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130811P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136312P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137503P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 18-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 22-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 23-JUN-1999; 99US-0140355P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153070P.
PR 15-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155589P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 3.0%; Score 14; DB 3; Length 336;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 PGKRVLDVCGGIGG 277
Db 94 PGKRVLDVCGGIGG 107

Search completed: July 26, 2004, 13:22:12
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:21:12 ; Search time 19 Seconds
(without alignments)
1285.215 Million cell updates/sec

Title: US-10-031-331B-40
Perfect score: 473
Sequence: 1 HTVDLTIAMMLDSQASLDL.....KLRRYAKGQRWGLFVAKKK 473

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:
1: /cgm2_6/ptodata/2/iaa/5A COMB.pdp:
2: /cgm2_6/ptodata/2/iaa/5B COMB.pdp:
3: /cgm2_6/ptodata/2/iaa/6A COMB.pdp:
4: /cgm2_6/ptodata/2/iaa/6B COMB.pdp:
5: /cgm2_6/ptodata/2/iaa/PCTUS COMB.pdp:
6: /cgm2_6/ptodata/2/iaa/backfiles1.pdp:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	2.3	363	3	US-09-041-718-5
2	10	2.1	318	4	US-09-382-906A-2
3	8	1.7	151	4	US-09-690-454-150
4	8	1.7	243	3	US-09-181-958-1
5	8	1.7	244	4	US-09-328-352-5622
6	7	1.7	248	4	US-09-543-681A-4328
7	8	1.7	266	4	US-09-489-039A-8384
8	8	1.7	348	4	US-09-118-637A-4
9	8	1.7	352	4	US-09-489-039A-9039
10	8	1.7	376	3	US-09-041-718-2
11	8	1.7	501	1	US-08-331-394-4
12	8	1.7	501	1	US-08-250-858-4
13	8	1.7	501	1	US-08-446-915-4
14	8	1.7	501	2	US-08-744-139-4
15	8	1.7	501	4	US-08-779-599-4
16	8	1.7	501	5	PCT-US95-06639-4
17	7	1.5	36	4	US-08-216-592A-18
18	7	1.5	179	4	US-09-134-001C-4112
19	7	1.5	220	4	US-09-823-823-22
20	7	1.5	245	4	US-09-205-258-369
21	7	1.5	273	4	US-09-252-991A-28278
22	7	1.5	298	4	US-09-644-907B-8
23	7	1.5	336	4	US-09-543-681A-4420
24	7	1.5	348	4	US-09-252-991A-29436
25	7	1.5	355	4	US-09-644-907B-4
26	7	1.5	379	4	US-09-134-000C-4846
27	7	1.5	383	3	US-09-041-718-3

ALIGNMENTS

RESULT 1
US-09-041-718-5
; Sequence 5, Application US/09041718A
; Patent No. 6225075
; GENERAL INFORMATION:
; APPLICANT: Bard, Martin
; TITLE OF INVENTION: DNA encoding sterol methyltransferase
; FILE REFERENCE: 740.003US1
; CURRENT APPLICATION NUMBER: US/09/041,718A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Triticum atrivum
US-09-041-718-5

Query Match 2.3%; Score 11; DB 3; Length 363;
Best Local Similarity 100.0%; Pred. No. 0.012; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 267 KVLVGGCGIGG 277
DB 122 KVLVGGCGIGG 132

RESULT 2
US-09-382-906A-2
; Sequence 2, Application US/09382906A
; Patent No. 6448475
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Shintani, David
; TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants
; FILE REFERENCE: 920905.90032
; CURRENT APPLICATION NUMBER: US/09/382,906A
; CURRENT FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 60/097,863
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 318
; ORGANISM: Synechocystis PCC6803
US-09-382-906A-2

Query Match 2.1%; Score 10; DB 4; Length 318;
Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VLDVGGIGG 277
|||||
Db 97 VLDVGGIGG 106

RESULT 3

US-09-690-454-150
; Sequence 150, Application US/09690454
; Patent No. 6531447

; GENERAL INFORMATION:

; APPLICANT: Steven M. Ruben, et al.

; TITLE OF INVENTION: 32 Human Secreted Proteins

; FILE REFERENCE: P2006PI

; CURRENT APPLICATION NUMBER: US/09/690,454

; CURRENT FILING DATE: 2000-10-18

; PRIOR APPLICATION NUMBER: 09/189,144

; PRIOR FILING DATE: 1998-11-10

; PRIOR APPLICATION NUMBER: 60/044,039

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,093

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,190

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/050,935

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,101

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/048,356

; PRIOR FILING DATE: May 30, 1997

; PRIOR APPLICATION NUMBER: 60/056,250

; PRIOR FILING DATE: August 29, 1997

; PRIOR APPLICATION NUMBER: 60/056,296

; PRIOR FILING DATE: August 29, 1997

; PRIOR APPLICATION NUMBER: 60/056,293

; PRIOR FILING DATE: August 29, 1997

; NUMBER OF SEQ ID NOS: 229

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 150

; LENGTH: 151

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (123)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-690-454-150

Query Match 1.7%; Score 8; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 SPTLSFPP 99
|||||
Db 47 SPTLSFPP 54

RESULT 4

US-09-181-958-1
; Sequence 1, Application US/09181958
; Patent No. 6143507

; GENERAL INFORMATION:

; APPLICANT: Kehry, Marilyn R

; APPLICANT: Pullen, Steven S

; APPLICANT: Crute, James J

; TITLE OF INVENTION: High Throughput Compatible Assays for Receptor-TRAF

; FILE REFERENCE: 9 142 Nucl. Seq

; CURRENT APPLICATION NUMBER: US/09/181,958

; CURRENT FILING DATE: 1998-10-29

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1
; LENGTH: 243
; TYPE: PRT
; ORGANISM: human
US-09-181-958-1

Query Match 1.7%; Score 8; DB 3; Length 243;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LRSIGLK 311
|||||
Db 53 LRSIGLK 60

RESULT 5

US-09-328-352-5622
; Sequence 5622, Application US/09328352
; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 5622

; LENGTH: 244

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-5622

Query Match 1.7%; Score 8; DB 4; Length 244;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
|||||
Db 60 KVLVDGCG 67

RESULT 6

US-09-543-681A-4328
; Sequence 4328, Application US/09543681A
; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 4328

; LENGTH: 248

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-4328

Query Match 1.7%; Score 8; DB 4; Length 248;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
|||||
Db 62 KVLVDGCG 69

RESULT 7

US-09-489-039A-8384


```
; Sequence 8384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8384
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8384

Query Match 1.7%; Score 8; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
Db 83 KVLVDVGGC 90

RESULT 8
US-09-118-637A-4
; Sequence 4, Application US/09118637A
; Patent No. 6642434
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Shintani, David K.
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/118,637A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 920905.90024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-118-637A-4

Query Match 1.7%; Score 8; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 DVGGGIGG 277
Db 270 DVGGGIGG 277

; Sequence 8384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8384
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8384

Query Match 1.7%; Score 8; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
Db 83 KVLVDVGGC 90

RESULT 8
US-09-118-637A-4
; Sequence 4, Application US/09118637A
; Patent No. 6642434
; GENERAL INFORMATION:
; APPLICANT: DellaPenna, Dean
; APPLICANT: Shintani, David K.
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/118,637A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 920905.90024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-118-637A-4

Query Match 1.7%; Score 8; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 DVGGGIGG 277
Db 270 DVGGGIGG 277

; Sequence 9399, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9039
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9039

Query Match 1.7%; Score 8; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
Db 209 KVLVDVGGC 216

RESULT 10
US-09-041-718-2
; Sequence 2, Application US/09041718A
; Patent No. 6225075
; GENERAL INFORMATION:
; APPLICANT: Bard, Martin
; TITLE OF INVENTION: DNA encoding sterol methyltransferase
; FILE REFERENCE: 740.003US1
; CURRENT APPLICATION NUMBER: US/09/041,718A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-041-718-2

Query Match 1.7%; Score 8; DB 3; Length 376;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGC 274
Db 125 KVLVDVGGC 132

RESULT 11
US-08-331-394-4
; Sequence 4, Application US/08331394
; Patent No. 5670319
; GENERAL INFORMATION:
; APPLICANT: Goeddel, David V.
; APPLICANT: Rothe, Mike
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-Associated Factors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
US-08-331-394-4

Query Match 1.7%; Score 8; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 DVGGGIGG 277
Db 270 DVGGGIGG 277
```

CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/331.394
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-331-394-4

Query Match 1.7%; Score 8; DB 1; Length 501;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311
DB 324 LERSIGLK 331

RESULT 12
US-08-250-858-4
Sequence 4, Application US/08250858
Patent No. 5708142
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/250.858
FILING DATE: 27-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897.1
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-250-858-4

Query Match 1.7%; Score 8; DB 1; Length 501;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311
DB 324 LERSIGLK 331

RESULT 13
US-08-446-915-4
Sequence 4, Application US/08446915
Patent No. 5741667
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.915
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/331394
FILING DATE: 28-OCT-1994
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 897P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-446-915-4

Query Match 1.7%; Score 8; DB 1; Length 501;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311
DB 324 LERSIGLK 331

RESULT 14
US-08-744-139-4
; Sequence 4, Application US/08779599
; Patent No. 5869612
; GENERAL INFORMATION:
; APPLICANT: Goedel, David V.
; APPLICANT: Rothe, Mike
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR-ASSOCIATED FACTORS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/744,139
FILING DATE: 31-Oct-1996
CLASSIFICATION: 435
PRIORITY APPLICATION NUMBER: 08/250858
FILING DATE: 05/27/1994
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0897C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-744-139-4

Query Match 1.7%; Score 8; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311
DB 324 LERSIGLK 331

RESULT 15
US-08-779-599-4
; Sequence 4, Application US/08779599
; Patent No. 6500922
; GENERAL INFORMATION:
; APPLICANT: Goedel, David V.
; APPLICANT: Rothe, Mike
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,599
FILING DATE: 07-Jan-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dregler, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P0897C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-779-599-4

Query Match 1.7%; Score 8; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 304 LERSIGLK 311
DB 324 LERSIGLK 331

Search completed: July 26, 2004, 13:24:28
Job time : 20 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:23:33 ; Search time 45 Seconds
(without alignments)
3291.599 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 473

Sequence: 1 HTVDLTIAMMLDSQASDLD.....KLRTAKGEQWGLFVAKK 473

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1288442 seqs, 313154207 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	5.5	219	12	US-10-425-114-53753
2	26	5.5	399	12	US-10-425-114-72568
3	26	5.5	513	12	US-10-425-114-37470
4	25	5.3	314	12	US-10-425-114-69972
5	23	4.9	331	12	US-10-425-114-44987
6	21	4.4	168	14	US-10-149-759-50
7	21	4.4	491	14	US-10-149-759-82
8	18	3.8	183	12	US-10-424-599-240673
9	17	3.6	133	16	US-10-437-963-110034
10	17	3.6	501	12	US-10-424-599-172601
11	17	3.6	530	16	US-10-437-963-155711
12	16	3.4	196	12	US-10-424-599-197735
13	16	3.4	210	12	US-10-424-599-197735
14	16	3.4	231	12	US-10-424-599-197735
15	16	3.4	246	12	US-10-424-599-197731

16	16	3.4	271	16	US-10-437-963-110035	Sequence 110035,
17	16	3.4	324	12	US-10-424-599-246283	Sequence 246283,
18	16	3.4	344	12	US-10-425-114-36573	Sequence 36573, A
19	16	3.4	344	12	US-10-425-114-36660	Sequence 36660, A
20	16	3.4	377	12	US-10-425-114-60338	Sequence 60338, A
21	16	3.4	408	12	US-10-424-599-197736	Sequence 197736,
22	14	3.0	120	16	US-10-437-963-155679	Sequence 155679,
23	13	2.7	237	12	US-10-424-599-257495	Sequence 257495,
24	11	2.3	58	12	US-10-424-599-201471	Sequence 201471,
25	11	2.3	138	16	US-10-437-963-167513	Sequence 167513,
26	11	2.3	202	10	US-09-768-235B-22	Sequence 22, Appl
27	11	2.3	285	15	US-10-259-134A-14	Sequence 14, Appl
28	11	2.3	338	12	US-10-425-114-41948	Sequence 41948, A
29	11	2.3	344	9	US-09-779-144A-7	Sequence 7, Appl
30	11	2.3	345	12	US-10-425-114-39343	Sequence 39343, A
31	11	2.3	345	16	US-10-437-963-147769	Sequence 147769, A
32	11	2.3	351	12	US-10-425-114-56209	Sequence 56209, A
33	11	2.3	351	12	US-10-425-114-65586	Sequence 65586, A
34	11	2.3	358	12	US-10-425-114-55779	Sequence 55779, A
35	11	2.3	358	12	US-10-425-114-59573	Sequence 59573, A
36	11	2.3	362	16	US-10-437-963-147768	Sequence 147768,
37	11	2.3	384	12	US-10-425-114-39121	Sequence 39121, A
38	11	2.3	383	16	US-10-437-963-184331	Sequence 184331,
39	11	2.3	409	16	US-10-437-963-167512	Sequence 167512,
40	11	2.3	412	12	US-10-425-114-42141	Sequence 42141, A
41	10	2.1	309	14	US-10-219-810-48	Sequence 48, Appl
42	10	2.1	318	14	US-10-219-810-46	Sequence 46, Appl
43	10	2.1	318	14	US-10-380-132-22	Sequence 22, Appl
44	10	2.1	318	16	US-10-471-243-20	Sequence 20, Appl
45	10	2.1	437	16	US-10-602-268-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-10-425-114-53753
; Sequence 53753, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53753
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-031-F4_FLI.pep
US-10-425-114-53753

Query Match 5.5% Score 26; DB 12; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.3e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 326 YPDNSFDVIVSRDILHIQDKPALFR 351

DB 72 YPDNSFDVIVSRDILHIQDKPALFR 97

RESULT 2

US-10-425-114-72568
; Sequence 72568, Application US/10425114
; Publication No. US20040034888A1

```
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 72566
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB23-061-D6_FLI.pep
US-10-425-114-72566

Query Match          5.5%; Score 26; DB 12; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.9e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDSFVYISRDITLHIQDKPALFR 351
DB 252 YPDSFVYISRDITLHIQDKPALFR 277

RESULT 3
US-10-425-114-37470
; Sequence 37470, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 37470
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB22-074-F3_FLI.pep
US-10-425-114-37470

Query Match          5.5%; Score 26; DB 12; Length 513;
Best Local Similarity 100.0%; Pred. No. 5e-17;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDSFVYISRDITLHIQDKPALFR 351
DB 366 YPDSFVYISRDITLHIQDKPALFR 391

RESULT 4
US-10-425-114-69972
; Sequence 69972, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
```

```
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 69972
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE108E03_FLI.pep
US-10-425-114-69972

Query Match          5.3%; Score 25; DB 12; Length 314;
Best Local Similarity 100.0%; Pred. No. 3.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MLDLKPQKVLVDGCGIGGDFYMA 283
DB 101 MLDLKPQKVLVDGCGIGGDFYMA 125

RESULT 5
US-10-425-114-44987
; Sequence 44987, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; CURRENT FILING DATE: 2003-04-28
; SEQ ID NO 44987
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700211781_FLI.pep
US-10-425-114-44987

Query Match          4.9%; Score 23; DB 12; Length 331;
Best Local Similarity 100.0%; Pred. No. 3.5e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLVDGCGIGGDFYMAE 284
DB 121 LKPGQKVLVDGCGIGGDFYMAE 143

RESULT 6
US-10-149-759-50
; Sequence 50, Application US/10149759
; Publication No. US20030157592A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
; TITLE OF INVENTION: Involved in the synthesis of tocopherols and
; FILE REFERENCE: BASF/NAE 1333/99 PCT/US
```

;; CURRENT APPLICATION NUMBER: US/10/149,759
;; CURRENT FILING DATE: 2002-10-17
;; PRIOR APPLICATION NUMBER: PCT/EP/00/12698
;; PRIOR FILING DATE: 2000-12-14
;; NUMBER OF SEQ ID NOS: 82
;; SOFTWARE: WordPerfect version 6.1
;; SEQ ID NO 50
;; LENGTH: 168
;; TYPE: PRT
;; ORGANISM: Physcomitrella patens
US-10-149-759-50

Query Match 4.4%; Score 21; DB 14; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 SFDVYSRDTLHIQDKPALF 350
DB 24 SFDVYSRDTLHIQDKPALF 44

RESULT 7

US-10-149-759-82
;; Sequence 82, Application US/10149759
;; Publication No. US20030157592A1
;; GENERAL INFORMATION:
;; APPLICANT: Lerchl, Jens
;; APPLICANT: Renz, Andreas
;; APPLICANT: Ehrhardt, Thomas
;; APPLICANT: Reindl, Andreas
;; APPLICANT: Cirpug, Petra
;; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
;; TITLE OF INVENTION: Involved in the synthesis of tocopherols and
;; TITLE OF INVENTION: carotenoids.
;; FILE REFERENCE: BASF/NAB 1333/99 PCT/US
;; CURRENT APPLICATION NUMBER: US/10/149,759
;; CURRENT FILING DATE: 2002-10-17
;; PRIOR APPLICATION NUMBER: PCT/EP/00/12698
;; PRIOR FILING DATE: 2000-12-14
;; NUMBER OF SEQ ID NOS: 82
;; SOFTWARE: WordPerfect version 6.1
;; SEQ ID NO 82
;; TYPE: PRT
;; ORGANISM: Physcomitrella patens
US-10-149-759-82

Query Match 4.4%; Score 21; DB 14; Length 491;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 PGGYIFFRSCPHQSGDHKK 152
DB 148 PGGYIFFRSCPHQSGDHKK 168

RESULT 8

US-10-424-599-240673
;; Sequence 240673, Application US/10424599
;; Publication No. US20040031072A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J
;; APPLICANT: Zhou Yihua
;; APPLICANT: Kovalic, David K
;; APPLICANT: Cao Yongwei
;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(53223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 285684
;; SEQ ID NO 240673
;; LENGTH: 183

;; TYPE: PRT
;; ORGANISM: Glycine max
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(183)
;; OTHER INFORMATION: unsure at all Xaa locations
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_59355C.1.pep
US-10-424-599-240673

Query Match 3.8%; Score 18; DB 12; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 SKCIGAYVRNKKQNOI 204
DB 96 SKCIGAYVRNKKQNOI 113

RESULT 9

US-10-437-963-110034
;; Sequence 110034, Application US/10437963
;; Publication No. US20040123343A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Wu, Wei
;; APPLICANT: Boukharov, Andrey A.
;; APPLICANT: Barbazuk, Brad
;; APPLICANT: Li, Ping
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(53221)B
;; CURRENT APPLICATION NUMBER: US/10/437,963
;; CURRENT FILING DATE: 2003-05-14
;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 110034
;; LENGTH: 133
;; TYPE: PRT
;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_14136C.1.pep
US-10-437-963-110034

Query Match 3.6%; Score 17; DB 16; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 EFAAYIKQGYDLHDVK 394
DB 39 EFAAYIKQGYDLHDVK 55

RESULT 10

US-10-424-599-172601
;; Sequence 172601, Application US/10424599
;; Publication No. US20040031072A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J
;; APPLICANT: Kovalic, David K
;; APPLICANT: Zhou Yihua
;; APPLICANT: Cao Yongwei
;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(53223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 285684
;; SEQ ID NO 172601
;; LENGTH: 501
;; TYPE: PRT

```
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (501)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_126875C.1.pep
US-10-424-599-172601

Query Match      3.6%; Score 17; DB 12; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 CKICGAYVRNKNQNOI 204
    |||||
DB 216 CKICGAYVRNKNQNOI 232
    |||||

RESULT 11
US-10-437-963-155711
; Sequence 155711, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Soukharov, Andrey A.
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155711
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (210)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5544C.1.pep
US-10-437-963-155711

Query Match      3.6%; Score 17; DB 16; Length 530;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 FQVYSRDTLHIOQKP 347
    |||||
DB 389 FQVYSRDTLHIOQKP 405
    |||||

RESULT 12
US-10-424-599-197735
; Sequence 197735, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 197735
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Glycine max
```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (196)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20580C.1.pep
US-10-424-599-197735

Query Match      3.4%; Score 16; DB 12; Length 196;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLGVCGGIGG 277
    |||||
DB 131 LKPGQKVLGVCGGIGG 146
    |||||

RESULT 13
US-10-424-599-246285
; Sequence 246285, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 246285
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) - (210)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_64427C.1.pep
US-10-424-599-246285

Query Match      3.4%; Score 16; DB 12; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLGVCGGIGG 277
    |||||
DB 92 LKPGQKVLGVCGGIGG 107
    |||||

RESULT 14
US-10-424-599-197733
; Sequence 197733, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 197733
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
```

```

; NAME/KEY: unsure
; LOCATION: (1).(231)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20579C.1.pep
US-10-424-599-197733

```

Query Match 3.4%; Score 16; DB 12; Length 231;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 262 LKPGQKVLDVCGGIGG 277
|||
db 141 LKPGQKVLDVCGGIGG 156

```

RESULT 15
US-10-424-599-197731
; Sequence 197731, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 197731
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(246)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20577C.1.pep
US-10-424-599-197731

```

```
Query Match      3.4%; Score 16; DB 12; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 262 LKPGQKVLDVCGIGG 277
|||
Db 138 LKPGQKVLDVCGIGG 153

Search completed: July 26, 2004, 13:29:12
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:20:17 ; Search time 17 Seconds
(without alignments)
2676.388 Million cell updates/sec

Title: US-10-031-331B-40
Perfect score: 473
Sequence: 1 HTVLTETMMQLDSQSLD.....KURRTAKGEQRWGLFVAKKK 473

Scoring table: OLIGO

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR-78.*

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	5.5	374	2	protein T1N15.23 (
2	26	5.5	555	2	hypothetical prote
3	16	3.4	367	2	probable sterol 24
4	12	2.5	346	2	sterol 24-C-methyl
5	11	2.3	344	2	sterol 24-C-methyl
6	11	2.3	344	2	sterol 24-C-methyl
7	11	2.3	363	2	probable sterol 24
8	10	2.1	318	2	hypothetical prote
9	9	1.9	280	2	gamma-cocopherol m
10	9	1.9	330	2	hypothetical prote
11	9	1.9	437	2	hypothetical prote
12	9	1.9	761	2	adenylate cyclase
13	9	1.9	1839	1	3-demethylubiquino
14	8	1.7	209	1	probable IS1560 tr
15	8	1.7	228	2	3-demethylubiquino
16	8	1.7	240	1	hypothetical prote
17	8	1.7	240	2	hypothetical prote
18	8	1.7	240	2	hypothetical prote
19	8	1.7	242	2	3-demethylubiquino
20	8	1.7	242	2	3-demethylubiquino
21	8	1.7	257	2	hypothetical prote
22	8	1.7	283	2	conserved hypothet
23	8	1.7	289	2	probable methyltra
24	8	1.7	342	2	conserved hypothet
25	8	1.7	343	2	ribosomal RNA smal
26	8	1.7	343	2	probable enzyme vj
27	8	1.7	343	2	tRNA (guanine-N2-)
28	8	1.7	348	2	gamma-cocopherol m
29	8	1.7	379	2	probable DELTA(24)

30 8 1.7 416 2 A70393 hypothetical prote
31 8 1.7 495 2 T27936 hypothetical prote
32 8 1.7 501 2 T61512 TNF receptor assoc
33 8 1.7 501 2 S56163 tumor necrosis fac
34 8 1.7 872 2 S49541 cellulase - Cellul
35 7 1.5 71 2 C69957 hypothetical prote
36 7 1.5 78 2 F82823 hypothetical prote
37 7 1.5 98 2 T15767 hypothetical prote
38 7 1.5 100 2 H71132 hypothetical prote
39 7 1.5 111 2 B45824 flal protein - Bac
40 7 1.5 111 2 B25159 13K sin operon hyp
41 7 1.5 128 1 FEHSX ferredoxin [2Fe-2S
42 7 1.5 128 2 S19497 hypothetical prote
43 7 1.5 129 1 FEHS ferredoxin [2Fe-2S
44 7 1.5 129 2 T43917 ferredoxin [2Fe-2S
45 7 1.5 129 2 B84380 ferredoxin import

ALIGNMENTS

RESULT 1

P96525 Protein T1N15.23 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: P96525
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwart, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: P96525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <STO>
A:Cross-references: GB:A8005173; MID:g8778697; PIDN:AAF79705.1; GSPDB:GN00141
C:Genetic: 1
A:Gene: T1N15.23
A:Map position: 1

Query Match 5.5%; Score 26; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 326 YDNSFDVYSRDTLHTQDKPALFR 351
DB 198 YDNSFDVYSRDTLHTQDKPALFR 223

RESULT 2

H96762 Hypothetical protein F6D5.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: H96762
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwart, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96762
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-555 <STO>
A:Cross-references: GB:AE005173; NID:g10092368; PIDN:AAG12776.1; GSPDB:GN00141
C:Genetics:
A:Gene: F6D5.1
A:Map position: 1

Query Match 5.5%; Score 26; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.4e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MLDLKPQKVLVDVCGIGSGDFYMAE 284

DB 341 MLDLKPQKVLVDVCGIGSGDFYMAE 366

RESULT 3
T06780
sterol 24-C-methyltransferase (SC 2.1.1.41) - soybean
N:Alternate names: S-adenosyl-L-methionine:delta24-sterol-C-methyltransferase
C:Species: Glycine max (soybean)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C:Accession: T06780
R:Shi, J.; Gonzalez, R.A.; Bhattacharyya, M.K.
J. Biol. Chem. 271, 9384-9389, 1996
A:Title: Identification and characterization of an S-adenosyl-L-methionine: delta 24-sterol
A:Reference number: Z15807; MUID:96199190; PMID:8621604
A:Accession: T06780
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-367 <SHI>
A:Cross-references: EMBL:U43683; NID:g1399379; PIDN:AAB04057.1; PID:g1399380
A:Experimental source: cultivar Williams 82; etiolated hypocotyls
C:Superfamily: 24-sterol C-methyltransferase; bioC homology
C:Keywords: methyltransferase; S-adenosylmethionine
F:122-226/Domain: bioC homology <BIOC>

Query Match 3.4%; Score 16; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 5.8e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 LKPGQKVLVDVCGIGG 277

DB 119 LKPGQKVLVDVCGIGG 134

RESULT 4
T10173
sterol 24-C-methyltransferase (SC 2.1.1.41) - castor bean
N:Alternate names: S-adenosyl-methionine-sterol-C-methyltransferase
C:Species: Ricinus communis (castor bean)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T10173
R:Bouvier-Nave, P.; Russelstein, T.; Desprez, T.; Benveniste, P.
Eur. J. Biochem. 246, 518-529, 1997
A:Title: Identification of cDNAs encoding sterol methyl-transferases involved in the sec
A:Reference number: Z15116; MUID:97352551; PMID:9208946
A:Accession: T10173
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-346 <BOU>
A:Cross-references: EMBL:U81313; NID:g2246457; PIDN:AAB62812.1; PID:g2246458
A:Experimental source: strain Baker 296
C:Superfamily: 24-sterol C-methyltransferase; bioC homology
C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 2.5%; Score 12; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 QKVLVDVCGIGG 277

DB 102 QKVLVDVCGIGG 113

RESULT 5
T01572
sterol 24-C-methyltransferase (EC 2.1.1.41) - maize
N:Alternate names: (S)-adenosyl-L-methionine:delta 24-sterol methyltransferase
C:Species: Zea mays (maize)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 03-Jun-2002
C:Accession: T01572
R:Tong, Y.; Nee, W.D.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z14350
A:Accession: T01572
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <TON>
A:Cross-references: EMBL:AF045570; NID:g2909845; PIDN:AAC04265.1; PID:g2909846
A:Experimental source: strain B73
C:Genetics:
A:Gene: SWI

C:Superfamily: 24-sterol C-methyltransferase; bioC homology
C:Keywords: methyltransferase; S-adenosylmethionine
F:101-205/Domain: bioC homology <BIOC>
Query Match 2.3%; Score 11; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVCGIGG 277

DB 103 KVLVDVCGIGG 113

RESULT 6
T04138
sterol 24-C-methyltransferase (EC 2.1.1.41) ESMT1, endosperm - maize
N:Alternate names: C-24 sterol methyltransferase
C:Species: Zea mays (maize)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C:Accession: T04138
R:Grebenok, R.J.; Galbraith, D.W.; Dellepenna, D.
Plant Mol. Biol. 34, 891-896, 1997
A:Title: Characterization of zeamays endosperm C-24 sterol methyltransferase - one of
A:Reference number: Z09668; MUID:97435974; PMID:9290641
A:Accession: T04138
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <GRE>
A:Cross-references: EMBL:U79669; NID:g1899059; PIDN:AAB70886.1; PID:g1899060
A:Experimental source: endosperm
C:Genetics:
A:Gene: ESMT1
C:Function:
A:Description: methyltransferase
C:Superfamily: 24-sterol C-methyltransferase; bioC homology
C:Keywords: methyltransferase; S-adenosylmethionine
F:101-205/Domain: bioC homology <BIOC>

Query Match 2.3%; Score 11; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVCGIGG 277

DB 103 KVLVDVCGIGG 113

RESULT 7
T06795
probable sterol 24-C-methyltransferase (EC 2.1.1.41) - wheat
N:Alternate names: delta-24-sterol methyltransferase

C:Species: Triticum aestivum (common wheat)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
 C:Accession: T06795
 R:Subramaniam, K.; Ueng, P.P.
 A:Reference number: 215820
 A:Reference number: 215820
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-363 <SUB>
 A:Cross-references: EMBL:U60754; NID:gl706964; PIDN:AB37769.1; PID:gl706965
 C:Genetics:
 A:Gene: MT1

C:Superfamily: 24-sterol C-methyltransferase; bioC homology
 C:Keywords: methyltransferase; S-adenosylmethionine
 F:120-224/Domain: bioC homology <BIOC>

Query Match 2.3%; Score 11; DB 2; Length 363;
 Best Local Similarity 100.0%; Pred. No. 0.008;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGGIGG 277
 |||||
 DB 122 KVLVGGGIGG 132

RESULT 8

S76226
 hypothetical protein - Synchocystis sp. (strain PCC 6803)

C:Species: Synchocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S76226
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis
 A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76226

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-318 <KAN>

A:Cross-references: EMBL:D90914; GB:AB001339; NID:gl653477; PIDN:BAAL8485.1; PID:dl01922

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: bioC homology

F:94-195/Domain: bioC homology <BIOC>

Query Match 2.1%; Score 10; DB 2; Length 318;
 Best Local Similarity 100.0%; Pred. No. 0.076;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VLDVGGGIGG 277
 |||||
 DB 97 VLDVGGGIGG 106

RESULT 9

AE2031
 gamma-tocopherol methyltransferase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AE2031
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2031

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-280 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073502.1; PID:gl7130893; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr1803

Query Match 1.9%; Score 9; DB 2; Length 280;
 Best Local Similarity 100.0%; Pred. No. 0.72;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 LDVGGGIGG 277
 |||||
 DB 67 LDVGGGIGG 75

RESULT 10

AC2071
 hypothetical protein all2121 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AC2071
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AC2071

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-330 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073820.1; PID:gl7131212; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all2121

Query Match 1.9%; Score 9; DB 2; Length 330;
 Best Local Similarity 100.0%; Pred. No. 0.83;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 LDVGGGIGG 277
 |||||
 DB 97 LDVGGGIGG 105

RESULT 11

T29330
 hypothetical protein F54D11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29330

R:Pauley, A.; Gattung, S.

A:Description: The sequence of C. elegans cosmid F54D11.

A:Reference number: Z20606

A:Accession: T29330

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-437 <PAU>

A:Cross-references: EMBL:U64834; PIDN:AB04824.1; GSPDB:GN00023; CESP:F54D11.1

A:Experimental source: strain Bristol N2; clone F54D11

C:Genetics:

A:Gene: CESP:F54D11.1

A:Map position: 5

A:Introns: 36/2; 76/3; 108/1; 315/3; 398/3

Query Match 1.9%; Score 9; DB 2; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 LKPGGKVLII 365
 |||||
 DB 318 LKPGGKVLII 326

RESULT 12

H65083
glycolate oxidase subunits GlcE and GlcF - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C/Accession: H65083
R/Blattner, P.R.; Maub, B.; Shao, Y.
A: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: H65083
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA

A/Residues: 1-761 <BLAT>
A/Cross-references: GB:A0000380; GB:U00096; NID:G1789344; PIDN:AAC76014.1; PID:G1789350;
A/Experimental source: strain K-12, substrain MG1655
C/Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hcd
P:372-447/Domain: ferredoxin 2[4Fe-4S] homology <FER2>

Query Match 1.9%; Score 9; DB 2; Length 761;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 VNEYQOMLK 401
|||||
DB 609 VNEYQOMLK 617

RESULT 13

OYBVK
adenylate cyclase (EC 4.6.1.1) - yeast (Saccharomyces kluyveri)
N/Alternate names: adenylyl cyclase
C/Species: Saccharomyces kluyveri
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 21-Jan-2000
C/Accession: JQ1145; S14464
R/Young, D.; O'Neill, K.; Broek, D.; Wigler, M.
Gene 102, 129-132, 1991
A/Title: The adenylyl cyclase-encoding gene from Saccharomyces kluyveri.
A/Reference number: JQ1145; MUID:91323718; PMID:1864503
A/Accession: JQ1145
A/Molecule type: DNA
A/Residues: 1-1839 <YOU>
A/Cross-references: EMBL:X56042; NID:G4856; PIDN:CAA39513.1; PID:G4857
C/Genetics:

A/Gene: CYR1
C/Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolog
C/Keywords: CAMP biosynthesis; duplication; phosphorus-oxygen lyase
P:1006-1029/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
P:1479-1564/Domain: yeast adenylate cyclase catalytic domain homology <YACC>

Query Match 1.9%; Score 9; DB 1; Length 1839;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LDFIESAIK 71
|||||
DB 650 LDFIESAIK 658

RESULT 14

S32628
3-demethylubiquinone-9 3-O-methyltransferase (EC 2.1.1.64) - Salmonella typhimurium (fra
N/Alternate names: 2-octaprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinone methyltran
C/Species: Salmonella typhimurium
C/Date: 28-May-1999 #sequence_revision 28-May-1999 #text_change 05-May-2000
C/Accession: S32628
R/Jordan, A.

submitted to the EMBL Data Library, March 1993

A/Reference number: S32628
A/Accession: S32628
A/Molecule type: DNA

A/Residues: 1-209 <JOR>

A/Cross-references: EMBL:X72948; NID:G510348; PIDN:CAA51451.1; PID:G295895
C/Superfamily: 3-demethylubiquinone-9 3-O-methyltransferase; bioC homology
C/Keywords: methyltransferase; S-adenosylmethionine; ubiquinone biosynthesis
P:24-126/Domain: bioC homology <BIOC>

Query Match 1.7%; Score 8; DB 1; Length 209;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVCGC 274
|||||
DB 26 KVLVDVCGC 33

RESULT 15

P70624
probable IS1560 truncated transposase Rv1035c - Mycobacterium tuberculosis (strain H37RV
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Sep-2000
C/Accession: F70624
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: F70624
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-228 <COL>

A/Cross-references: GB:Z92539; GB:AL123456; NID:G3261714; PIDN:CAB06844.1; PID:e304621;
A/Experimental source: strain H37RV
C/Genetics:

A/Gene: Rv1035c

Query Match 1.7%; Score 8; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 RPTGELAE 55
|||||
DB 121 RPTGELAE 128

Search completed: July 26, 2004, 13:23:58

Job time: 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:16:27 ; Search time 13 Seconds
(without alignments)
1894.553 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 473

Sequence: 1 HTVDTLTETAMLDQASOLD.....KLRTAKGEQWGLFVAKKK 473

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	9.5	494	PEAM_SPIOL	Q9m571 spinacia ol
2	26	5.5	475	PEM3_ARATH	Q944h0 arabidopsis
3	26	5.5	490	PEM3_ARATH	Q946b9 arabidopsis
4	26	5.5	491	PEM3_ARATH	Q946b9 arabidopsis
5	9	1.9	365	1D12_METKA	Q8x999 methanopyru
6	9	1.9	407	1D12_METKA	P52074 escherichia
7	9	1.9	1839	1CYAA_SACKL	P23466 saccharomyc
8	8	1.7	232	1UBIG_CHRVO	Q7nz91 chromobacte
9	8	1.7	232	1UBIG_PSEPK	Q88m10 pseudomonas
10	8	1.7	232	1UBIG_PSEPK	Q88m10 pseudomonas
11	8	1.7	235	1UBIG_VIRPA	Q87nd5 vibrio para
12	8	1.7	235	1UBIG_VIRVU	Q88e0 vibrio vuln
13	8	1.7	235	1UBIG_VIRVU	Q7mm27 vibrio vuln
14	8	1.7	236	1UBIG_HABDU	Q7vkw2 haemophilus
15	8	1.7	236	1UBIG_SHEON	Q8ee99 shewanella
16	8	1.7	240	1UBIG_ECOL6	Q8xe29 escherichia
17	8	1.7	240	1UBIG_ECOL6	Q8ffp0 escherichia
18	8	1.7	240	1UBIG_ECOL1	P17993 escherichia
19	8	1.7	240	1UBIG_PHOLL	Q7n2m5 photorhabd
20	8	1.7	240	1UBIG_SHIFL	Q820c5 shigella fl
21	8	1.7	242	1UBIG_PASMU	Q9cm16 pasteurella
22	8	1.7	242	1UBIG_SALTI	Q8z560 salmonella
23	8	1.7	242	1UBIG_SALTI	P37431 salmonella
24	8	1.7	242	1UBIG_YERPE	Q8z9re yersinia pe
25	8	1.7	342	1RSMC_ECOLI	P39406 escherichia
26	8	1.7	348	1GTOM_ARATH	Q9z8k1 arabidopsis
27	8	1.7	376	1BRG6_CANAL	Q14198 candida alb
28	8	1.7	501	1TRA2_HUMAN	Q12933 homo sapien
29	8	1.7	501	1TRA2_MOUSE	P39429 mus musculu
30	8	1.7	572	1GUXA_CELFI	P50401 cellulomona
31	7	1.5	71	1YQOQ_BACSU	P54494 bacillus su
32	7	1.5	111	1SINR_BACLI	P22753 bacillus li
33	7	1.5	111	1SINR_BACSU	P06533 bacillus su

RESULT 1

ID	PEAM_SPIOL	STANDARD;	PRT;	494 AA.
AC	Q9M571;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Phosphoethanolamine N-methyltransferase (EC 2.1.1.103).			
GN	PEMT.			
OS	Spinacia oleracea (Spinach).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Caryophyllales; Amaranthaceae; Spinacia.			
OX	NCBI_TaxID=3562;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	STRAIN=cv. Savoy hybrid 612; TISSUE=Leaf.			
RX	MEDLINE=20261526; PubMed=10799484;			
RA	Nuccio M.L., Ziemak M.J., Henry S.A., Weretilnyk E.A., Hanson A.D.;			
RT	"Phosphoethanolamine N-methyltransferase from spinach: cDNA cloning by			
RT	complementation in Schizosaccharomyces pombe and characterization of			
RT	the recombinant enzyme."			
RL	J. Biol. Chem. 275:14095-14101(2000).			
CC	FUNCTION: Catalyzes N-methylation of phosphoethanolamine,			
CC	phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the			
CC	three methylation steps required to convert phosphoethanolamine to			
CC	phosphocholine. Mediates a key step in the biosynthesis of			
CC	choline, a precursor of the osmoprotectant glycine betaine. Has an			
CC	optimal pH of 7.8 to 8.5. Has no ethanolamine- or			
CC	phosphatidylethanolamine-N-methyltransferase activity.			
CC	CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine			
CC	phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine			
CC	phosphate.			
CC	ENZYME REGULATION: Inhibited by phosphocholine but not by choline,			
CC	glycine betaine, monomethyl ethanolamine or dimethylethanolamine.			
CC	SUBUNIT: Monomer.			
CC	SUBCELLULAR LOCATION: Cytoplasmic.			
CC	INDUCTION: By salt stress.			
CC	SIMILARITY: Belongs to the methyltransferase superfamily.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AF237623; AAP61950.1;			
DR	InterPro; IPR001601; Methyltransf.			
DR	InterPro; IPR000051; SAM_bind.			
DR	Methyltransferase; Transferase; Repeat.			
KM	DOMAIN 53...162			
FT	DOMAIN 282...389			
FT	DOMAIN 494 AA; 56361 MW; 7F2537C8E4B8413B CRC64;			
SQ	SEQUENCE			

P00217 haloarcula
P00216 halobacteri
P25649 saccharomyc
O67445 aquifex aeo
Q8rdm9 fuscobacteri
Q10886 mycobacteri
Q9c3n8 pasteurella
P77889 lactobacill
Q35423 synechocyst
Q9c3a8 mycobacteri
O06424 mycobacteri
Q82055 coxiella bu

ALIGNMENTS

```

Query Match          9.5%; Score 45; DB 1; Length 494;
Best Local Similarity 100.0%; Pred. No. 2.2e-37;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 LVERMLKXKGGYVIFRSCFQSGDHKRSNTHYRPRVTK 166
DB 143 LVERMLKXKGGYVIFRSCFQSGDHKRSNTHYRPRVTK 187

RESULT 2
ID PEM2 ARATH STANDARD; PRT; 475 AA.
AC 0944H0; Q9LP63; Q9LP64;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Putative phosphoethanolamine N-methyltransferase 2 (EC 2.1.1.103).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Cressey T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Luros J.S., Maiti R., Marziani A.,
RA Lin X., Liu S.X., Li Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzgier S.L., Tambunga G., Toriumi M.J., Town C.D.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RA thaliana."
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayaishi Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Kani S., Koesema G., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Khan S., Koesema G., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kani S., Koesema G., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Satou M., Tamura E., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Tsuruta S., Shinohara K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome."
RL Science 302:842-846(2003).
CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,
CC phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the
CC three methylation steps required to convert phosphoethanolamine to
CC phosphocholine (By similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine
CC phosphate = S-adenosyl-L-homocysteine + N-methyl ethanolamine
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

```

```

CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.
CC -1- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AC020889; AAF79704.1; ALT_SEQ.
CC EMBL; AC020889; AAF79705.1; ALT_SEQ.
CC EMBL; AF428454; AAL16223.1;
CC EMBL; AY063866; AAL36222.1;
CC InterPro; IPR001601; Methyltransf.
CC InterPro; IPR000051; SAM_Bind.
CC Pfam; PF01209; UbiE/methyltransf. 1.
CC Transferrase; Methyltransferase; Repeat.
CC TRANSFAM; 34 143
CC FT DOMAIN 263 370 SAM-BINDING 2.
CC SEQUENCE 475 AA; 54018 MW; 040705A4DE80DA8 CRC64;
CC
Query Match          5.5%; Score 26; DB 1; Length 475;
Best Local Similarity 100.0%; Pred. No. 3.9e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDNSFDVYSRDTILHIQDKPALFR 351
DB 328 YPDNSFDVYSRDTILHIQDKPALFR 353

RESULT 3
ID PEM3 ARATH STANDARD; PRT; 490 AA.
AC 09C8E9; Q9C9V1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Putative phosphoethanolamine N-methyltransferase 3 (EC 2.1.1.103).
GN NMT3 OR AT1G73600 OR F6D5.1 OR F25P22.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Cressey T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Luros J.S., Maiti R., Marziani A.,
RA Lin X., Liu S.X., Li Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzgier S.L., Tambunga G., Toriumi M.J., Town C.D.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RA thaliana."
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Cressey T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Luros J.S., Maiti R., Marziani A.,
RA Lin X., Liu S.X., Li Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzgier S.L., Tambunga G., Toriumi M.J., Town C.D.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RA thaliana."
RL Nature 408:816-820(2000).
RN [2]
RP CONCEPTUAL TRANSLATION.

```

RA Schneider M.;
 CC Unpublished observations (MAY-2002).
 CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine, the
 CC phosphononmethylethanolamine and phosphodimethylethanolamine, the
 CC three methylation steps required to convert phosphoethanolamine to
 CC phosphocholine (By similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine
 CC phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.
 CC -1- CAUTION: Ref.1 (ARGS1806) sequence differs from that shown due to
 CC erroneous gene model prediction.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AC079676; ARGS1806.1; ALT_SEQ.
 CC DR EMBL; AC012679; ARGS2075.1; -
 CC DR InterPro; IPR001601; Methyltransf.
 CC DR InterPro; IPR000051; SAM bind.
 CC KW Transferrase; Methyltransferase; Repeat.
 CC FT DOMAIN 49 158 SAM-BINDING 1.
 CC FT DOMAIN 278 385 SAM-BINDING 2.
 CC FT SEQUENCE 490 AA; 56368 MW; 77DFAFBC9C41CB CRC64;
 CC
 CC Query Match 5.5%; Score 26; DB 1; Length 490;
 CC Best Local Similarity 100.0%; Pred. No. 4e-18; 0; Indels 0; Gaps 0;
 CC Matches 26; Conservative 0; Mismatches 0;
 CC
 CC QY 259 MLDLXPGQKVLVDVCGIGGGDFYNAE 284
 CC DB 276 MLDLXPGQKVLVDVCGIGGGDFYNAE 301
 CC
 CC RESULT 4
 CC FEMI ARATH STANDARD; PRT; 491 AA.
 CC ID FEMI ARATH STANDARD; PRT; 491 AA.
 CC AC Q9FR44; O9LVH3; -
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Phosphoethanolamine N-methyltransferase 1 (EC 2.1.1.103) (PEAMT 1)
 CC (AtNMT1).
 CC GN NMT1 OR AT3G17990/AT3G18000 OR MEBS.19 OR MEBS.21/MEBS.22.
 CC OS Arabidopsis thaliana (Mouse-ear cress).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 CC OX NCBI_TaxID=3702;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Landsberg erecta; TISSUE=Seedling;
 CC RX MEDLINE=20567827; PubMed=11115895;
 CC RA Bologna C.P., McGraw P.;
 CC RT "The isolation and characterization in yeast of a gene for Arabidopsis
 CC S-adenosylmethionine:phospho-ethanolamine N-methyltransferase.";
 CC RL Plant Physiol. 124:1800-1813 (2000).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Columbia;
 CC RX MEDLINE=2077480; PubMed=10819129;
 CC RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
 CC RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 CC features of the regions of 4,504,864 bp covered by sixty pl and PAC
 CC clones.";
 CC RL DNA Res. 7:131-135 (2000).
 CC RN [3]

RP SEQUENCE FROM N.A.
 CC STRAIN=cv. Columbia;
 CC RX MEDLINE=22954850; PubMed=14591372;
 CC RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 CC Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.P.,
 CC Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 CC Miranda M., Quach H.B., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 CC Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 CC Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 CC Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjail M., Hansen M.F.,
 CC Hayashizaki Y., Johnson-Hopson C., Heuan V.W., Iida K., Karnes M.,
 CC Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 CC Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 CC Satou M., Tamse R., Vayberg M., Wallender E.K., Wong C., Yamamura Y.,
 CC Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 CC RT "Empirical analysis of transcriptional activity in the Arabidopsis
 CC genome.";
 CC RL Science 302:842-846 (2003).
 CC CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,
 CC phosphononmethylethanolamine and phosphodimethylethanolamine, the
 CC three methylation steps required to convert phosphoethanolamine to
 CC phosphocholine.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine
 CC phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.
 CC -1- CAUTION: This protein is a fusion of the two annotated genes
 CC AT3G17990 and AT3G18000.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF197940; ARG41121.1; -
 CC DR EMBL; AB019230; BAB02720.1; ALT_SEQ.
 CC DR EMBL; AF367299; AAK32886.1; -
 CC DR EMBL; AY058175; AAL25589.1; -
 CC DR EMBL; AY091683; AAM10282.1; -
 CC DR EMBL; AY093093; AAM13092.1; -
 CC DR InterPro; IPR001601; Methyltransf.
 CC DR InterPro; IPR000051; SAM bind.
 CC KW Methyltransferase; Transferase; Repeat.
 CC FT DOMAIN 50 159 SAM-BINDING 1.
 CC FT DOMAIN 279 386 SAM-BINDING 2.
 CC FT CONFLICT 333 333 E -> G (IN REF. 3; AAM13092).
 CC FT SEQUENCE 491 AA; 56102 MW; DA4404BED3FA8D5 CRC64;
 CC
 CC Query Match 5.5%; Score 26; DB 1; Length 491;
 CC Best Local Similarity 100.0%; Pred. No. 4e-18;
 CC Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 326 YPDNSFDVIYSRDTILHIQDKPALFR 351
 CC DB 344 YPDNSFDVIYSRDTILHIQDKPALFR 369
 CC
 CC RESULT 5
 CC ID12 METKA STANDARD; PRT; 365 AA.
 CC AC Q9TX99;
 CC DT 10-OCT-2003 (Rel. 42, Created)
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2) (IPP isomerase)
 CC DE (Isopentenyl pyrophosphate isomerase).
 CC GN FNI OR LDD OR MK0776.

OS Methanopyrus kandleri.
OC Archaea: Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Shestarev A.I., Mezheva V.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Kozlov E.V., Koryavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
CC -1- FUNCTION: Catalyzes the 1,3-allylic rearrangement of the
CC homoallylic substrate isopentenyl (IPP) to its allylic isomer,
CC dimethylallyl diphosphate (DMAPP) (By similarity).
CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate = dimethylallyl
CC diphosphate.
CC -1- COFACTOR: FMN and NADPH (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the IPP isomerase type 2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF010369; AAM01990.1; ALT_INIT.
DR HAMAP; MF_00354; -; 1.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR00262; FMN_hydryac_dh.
DR Pfam; PF01070; FMN_dh; 1.
DR Isomerase; Isoprene biosynthesis; Flavoprotein; FMN; NADP;
KW Complete proteome.
SQ SEQUENCE 365 AA; 39272 MW; B5C0541EBB49C355 CRC64;

Query Match 1.9%; Score 9; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 403 AGFVDVLAE 411
DB 167 AGFVDVLAE 175

RESULT 6
GLCF_ECOLI
ID GLCF_ECOLI STANDARD; PRT; 407 AA.
AC P52074; P76654;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glycolate oxidase iron-sulfur subunit.
DE GLCF OR GOX R 82978.
OS Escherichia coli
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RX MEDLINE=96178972; PubMed=8606183;
RA Pellicer M.T., Badia J., Aguilar J.T., Baldoma L.,
RA "The glc locus of Escherichia coli: characterization of genes encoding
RT the subunits of glycolate oxidase and the glc regulator protein";
RL J. Bacteriol. 178:2051-2059 (1996).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-K12 / M31655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blactner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1234-1238 (1997).
CC -1- SIMILARITY: The iron-sulfur centers are similar to those of
CC bacterial-type 4Fe-4S ferredoxins.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT THAT FUSES TOGETHER GLCE AND GLCF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L43490; AAB02532.1; -;
DR EMBL; U28377; AAB69145.1; ALT_FRAME.
DR EMBL; AE000380; AAC76014.1; ALT_FRAME.
DR EcoGene; EG13291; glcf.
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR InterPro; IPR004017; DUF224.
DR Pfam; PF02754; CCG; 2.
DR Pfam; PF00033; fer4; 2.
DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 2.
KW Electron transport; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 25 25 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 28 28 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 31 31 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 35 35 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 75 75 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 78 78 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 81 81 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 85 85 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 407 AA; 45110 MW; F983C893063F76DC CRC64;

Query Match 1.9%; Score 9; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 VKKEYGQMLK 401
DB 255 VKKEYGQMLK 263

RESULT 7
CYAA_SACKL
ID CYAA_SACKL STANDARD; PRT; 1839 AA.
AC P23466;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylate
DE cyclase).
DE CYR.
OS Saccharomyces kluyveri (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4934;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=9132718; PubMed=1864503;
RX Young D., O'Neill K., Broek D., Wigler M.;
RT "The adenylate cyclase-encoding gene from Saccharomyces kluyveri";
RL Gene 102:129-132 (1991).
CC -1- FUNCTION: Plays essential roles in regulation of cellular
CC metabolism by catalyzing the synthesis of a second messenger,
CC cAMP.

UBIG_PSBPK STANDARD; PRT; 232 AA.

AC Q88M10;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHBB
 methyltransferase).
 GN UBIG OR PFI1765.

OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weigel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Neeson M., Deboy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson M., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RA "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AS016780; AAN67385.1; -
 CC TIGR; PFI1765; -
 CC HAMAP; MF_00472; -; 1.
 CC InterPro; IPR001601; Methyltransf.
 CC InterPro; IPR000051; SAM bind.
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.
 CC Pfam; PF01209; Ubie_methyltran; 1.
 CC TIGRFAMs; TIGR01983; UbiG; 1.
 CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
 CC Complete proteome.
 CC SEQUENCE 232 AA; 26084 MW; 7A795B7DF50479B9 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 232;
 Best Local Similarity 100.0%; Pred.No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
 DB 50 KVLVDGCG 57
 |||||
 |||||

RESULT 10
 UBIG_PSESM STANDARD; PRT; 232 AA.

AC Q885T9;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHBB
 methyltransferase).
 GN UBIG OR PFI1765.

OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weigel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Neeson M., Deboy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson M., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RA "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AS016780; AAN67385.1; -
 CC TIGR; PFI1765; -
 CC HAMAP; MF_00472; -; 1.
 CC InterPro; IPR001601; Methyltransf.
 CC InterPro; IPR000051; SAM bind.
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.
 CC Pfam; PF01209; Ubie_methyltran; 1.
 CC TIGRFAMs; TIGR01983; UbiG; 1.
 CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
 CC Complete proteome.
 CC SEQUENCE 232 AA; 26084 MW; 7A795B7DF50479B9 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 232;
 Best Local Similarity 100.0%; Pred.No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
 DB 50 KVLVDGCG 57
 |||||
 |||||

RESULT 11
 UBIG_VIBPA STANDARD; PRT; 235 AA.

AC Q87ND5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 3-demethylubiquinone-9 3-methyltransferase (SC 2.1.1.64) (3,4-
 dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHBB
 methyltransferase).
 GN UBIG OR VPI933.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OC NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=222834015; PubMed=12928499;
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
 RA Quinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
 RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
 RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,
 RA Khouri H., Fedorova N., Tran B., Russell D., Berty K., Utterback T.,
 RA Van Aken S.E., Feldblum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
 RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,
 RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
 RA White O., Fraser C.M., Collier A.;
 RA "The complete genome sequence of the Arabidopsis and tomato pathogen
 RT Pseudomonas syringae pv. tomato DC3000.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE016862; AOS55262.1; ALT_INIT.
 CC TIGR; PSPT01742; -; 1.
 CC HAMAP; MF_00472; -; 1.
 CC InterPro; IPR001601; Methyltransf.
 CC InterPro; IPR000051; SAM bind.
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.
 CC Pfam; PF01209; Ubie_methyltran; 1.
 CC TIGRFAMs; TIGR01983; UbiG; 1.
 CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
 CC Complete proteome.
 CC SEQUENCE 232 AA; 25932 MW; 248F3D231401A9E4 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 232;
 Best Local Similarity 100.0%; Pred.No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDGCG 274
 DB 50 KVLVDGCG 57
 |||||
 |||||

```

EX MEDLINE=22509454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nishijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yaenaga T., Honda T., Shinagawa H., Hattori M., Iida T.,
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005079; BAC60136.1; -.
DR HAMAP; MF_00472; -.
DR InterPro; IPR001601; Methyltransf.
DR Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26179 MW; 934D8E4403D017F0 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGCG 274
DB 54 KVLVGGCG 61

RESULT 12
UBIG_VIBVU STANDARD; PRT; 235 AA.
AC QD8E0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
DE methyltransferase).
GN UBIG OR VV13040.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen C.Y., Wu K.M., Chang Y.C., Chang C.H., Tsai H.C., Liao T.L.,
RA Liu Y.M., Chen H.J., Shen A.B., Li J.C., Su T.L., Shao C.P., Lee C.T.,
RA Hor L.I., Tsai S.F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005335; BAC94010.1; -.
DR HAMAP; MF_00472; -.
DR Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26244 MW; A89B7E1F8613F2B8 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGCG 274
DB 54 KVLVGGCG 61

RESULT 14
UBIG_HABDU STANDARD; PRT; 236 AA.
AC Q7VKW2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```

```

DR HAMAP; MF_00472; -.
DR InterPro; IPR001601; Methyltransf.
DR Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26244 MW; A89B7E1F8613F2B8 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGCG 274
DB 54 KVLVGGCG 61

RESULT 13
UBIG_VIBVU STANDARD; PRT; 235 AA.
AC Q7NM27;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
DE methyltransferase).
GN UBIG OR VV1246.
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen C.Y., Wu K.M., Chang Y.C., Chang C.H., Tsai H.C., Liao T.L.,
RA Liu Y.M., Chen H.J., Shen A.B., Li J.C., Su T.L., Shao C.P., Lee C.T.,
RA Hor L.I., Tsai S.F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005335; BAC94010.1; -.
DR HAMAP; MF_00472; -.
DR Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 235 AA; 26244 MW; A89B7E1F8613F2B8 CRC64;

Query Match 1.7%; Score 8; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVGGCG 274
DB 54 KVLVGGCG 61

RESULT 14
UBIG_HABDU STANDARD; PRT; 236 AA.
AC Q7VKW2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```

DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
 DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHS
 DE UBIG OR HD1750)
 GN Haemophilus ducreyi.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3500HP / ATCC 700724;
 RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
 RA "The complete genome sequence of Haemophilus ducreyi.";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AB017155; AAP96505.1; -;
 DR HAMAP; MF 00472; -; 1.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 DR TIGRFAWS; TIGR01983; UbiG; 1.
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 236 AA; 26602 MW; B266FC407FF6AECDC CRC64;
 Query Match 1.7%; Score 8; DB 1; Length 236;
 Best Local Similarity 100.0%; Pred. No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 267 KVLVDGCG 274
 DB 51 KVLVDGCG 58
 RESULT 15
 UBIG SHEON STANDARD; PRT; 236 AA.
 AC QSEEG;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
 DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHHS
 DE UBIG OR SO2413)
 GN Shewanella oneidensis.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 OX NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NR-1;
 RX MEDLINE=22297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupati K., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,

RA Feldblyum T.V., Smith H.O., Venter J.C., Nealeson K.H., Fraser C.M.;
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 RL *Shewanella oneidensis*.";
 RL Nat. Biotechnol. 20:1118-1123(2002).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ubiG/COQ3 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AB015682; AAN55447.1; -;
 DR TIGR; SO2413; -; 1.
 DR HAMAP; MF 00472; -; 1.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 DR TIGRFAWS; TIGR01983; UbiG; 1.
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
 KW Complete proteome.
 SQ SEQUENCE 236 AA; 26469 MW; A4F0C932B5533E76 CRC64;
 Query Match 1.7%; Score 8; DB 1; Length 236;
 Best Local Similarity 100.0%; Pred. No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 267 KVLVDGCG 274
 DB 54 KVLVDGCG 61
 Search completed: July 26, 2004, 13:22:37
 Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:19:37 ; Search time 42 Seconds
(Without alignments)
3553.336 Million cell updates/sec

Title: US-10-031-331B-40
Perfect score: 473
Sequence: 1 HTVDLTIEMMLDSQASDLID.....KLRRTAKGEQWGLFVAKKK 473

Scoring table: OUIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	473	100.0	494	10 Q852S7	Q852S7 suaeda japo
2	27	5.7	491	10 Q9AXH3	Q9AXH3 lycopersico
3	26	5.5	376	10 Q8L7A8	Q8L7A8 arabidopsis
4	26	5.5	493	10 Q846A4	Q846A4 aster tripo
5	24	5.1	491	10 Q7XJ22	Q7XJ22 braesica na
6	21	4.4	498	10 Q8VYX1	Q8VYX1 triticum ac
7	17	3.6	499	10 Q8LJ10	Q8LJ10 oryza sativ
8	16	3.4	346	10 Q82720	Q82720 nicotiana t
9	16	3.4	349	10 Q82434	Q82434 nicotiana t
10	16	3.4	367	10 Q43445	Q43445 glycine max
11	14	3.0	336	10 Q8LKW1	Q8LKW1 arabidopsis
12	14	3.0	336	10 Q9LM02	Q9LM02 arabidopsis
13	12	2.5	346	10 Q24328	Q24328 ricinus com
14	11	2.3	330	10 Q84M50	Q84M50 oryza sativ
15	11	2.3	344	10 P93852	P93852 sea maye (m
16	11	2.3	344	10 Q49215	Q49215 sea maye (m

17	11	2.3	349	10 Q82426	Q82426 oryza sativ
18	11	2.3	363	10 Q41587	Q41587 triticum ae
19	11	2.3	363	10 Q41586	Q41586 triticum ae
20	10	2.1	87	16 Q7VD26	Q7VD26 prochloroco
21	10	2.1	304	16 Q7V427	Q7V427 prochloroco
22	10	2.1	309	16 Q7U4C8	Q7U4C8 synchococ
23	10	2.1	318	16 P74388	P74388 synchococ
24	10	2.1	328	16 Q8DI65	Q8DI65 synchococ
25	9	1.9	247	16 Q882P3	Q882P3 rhizobium l
26	9	1.9	255	16 Q81TP8	Q81TP8 bacillus an
27	9	1.9	273	2 Q8KI52	Q8KI52 nocardia ae
28	9	1.9	280	16 Q8YWL3	Q8YWL3 anabaena sp
29	9	1.9	283	2 Q8KZ94	Q8KZ94 nocardia ae
30	9	1.9	311	16 Q7UZZ0	Q7UZZ0 prochloroco
31	9	1.9	330	16 Q8YV60	Q8YV60 anabaena sp
32	9	1.9	341	5 Q86I59	Q86I59 dictyosteli
33	9	1.9	437	5 Q22993	Q22993 caenorhabdi
34	9	1.9	2094	11 Q80Y35	Q80Y35 mus musculu
35	8	1.7	88	17 Q970B1	Q970B1 sulfolobus
36	8	1.7	202	17 Q8PUL7	Q8PUL7 methanosarc
37	8	1.7	207	16 Q8XNS1	Q8XNS1 clostridium
38	8	1.7	221	16 Q8F5S7	Q8F5S7 leptospira
39	8	1.7	223	16 Q8YUL6	Q8YUL6 bradyrhizob
40	8	1.7	228	16 P96366	P96366 mycobacteri
41	8	1.7	228	16 Q7UUI7	Q7UUI7 mycobacteri
42	8	1.7	229	17 Q8TRV4	Q8TRV4 methanosarc
43	8	1.7	232	16 Q88M10	Q88M10 pseudomonas
44	8	1.7	236	16 Q8EEG9	Q8EEG9 shewanella
45	8	1.7	236	16 Q7VKW2	Q7VKW2 haemophilus

ALIGNMENTS

RESULT 1

Q852S7 PRELIMINARY; PRT; 494 AA.

AC Q852S7; DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Phosphoethanolamine N-methyltransferase.

GN PEAMT.

OS Sueda japonica.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;

OC Caryophyllales; Amaranthaceae; Sueda.

OX NCBI_TaxID=90346;

RP SEQUENCE FROM N.A.

RT Yamada A., Nozawa G.T., Tanimoto S., Ozeki Y.;

RL Submitted (F88-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB080186; BAC57432.1; -

DR GO; GO:0008757; F:8-adenosylmethionine-dependent methyltransf. . . ; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR InterPro; IPR001601; Methyltransf.

DR InterPro; IPR000051; SAM_Bind.

DR InterPro; IPR004033; UbiE/COQ5_Metrf.

DR Pfam; PF01209; UbiE_methyltran; 1.

KW Transferase; Methyltransferase.

SQ SEQUENCE 494 AA; 56562 MW; BC613F9097BD3AE3 CRC64;

Query Match 100.0%; Score 473; DB 10; Length 494;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HTVDLTIEMMLDSQASDLIDKERPERILSNLPLPBGKCLLELGACIGFTGELAFKAGQV 60

DB 27 HTVDLTIEMMLDSQASDLIDKERPERILSNLPLPBGKCLLELGACIGFTGELAFKAGQV 81

QY 61 TALDFIESAIKKEVINGHYKNVKNFCADVTSPTLSFPFHSIDVIFSNLWMLYLSDEVE 120

Db 82 IALDPIESAIKNEVINGHYKNKWCADVTPTLSFPFPHSLDVFISNMLLMVLSDEVE 141
QY 121 NLVERMLKMLKPGYIFPESCHQSGHRSNTHVREPRFYTKAFKCHLODGSNS 180
Db 142 NLVERMLKMLKPGYIFPESCHQSGHRSNTHVREPRFYTKAFKCHLODGSNS 201
QY 181 YELSLSCKICIGAVVRNKNQNIQSWLWQVDSKDDKQFQPLDTSQYKCNILRYERV 240
Db 202 YELSLSCKICIGAVVRNKNQNIQSWLWQVDSKDDKQFQPLDTSQYKCNILRYERV 261
QY 241 GPGVSTGGVETTKFVSMMLDKPGQKVDVGGGDFYMAETFDVVGVGFLSVNMI 300
Db 262 GPGVSTGGVETTKFVSMMLDKPGQKVDVGGGDFYMAETFDVVGVGFLSVNMI 321
QY 301 SPALERSLGLKCAVEPAVADCTKINPDNSFDVYSRDTILHIQDKPALFRSFKMLKPG 360
Db 322 SPALERSLGLKCAVEPAVADCTKINPDNSFDVYSRDTILHIQDKPALFRSFKMLKPG 381
QY 361 GRVLISYCKKAGPPSPFAAYTKQGYDLHVKYQMLKQDAGFVVLADRTQFIRV 420
Db 382 GRVLISYCKKAGPPSPFAAYTKQGYDLHVKYQMLKQDAGFVVLADRTQFIRV 441
QY 421 LRKLETVKEKDVIFISDFSEEDYNDIVGMDKLRATKAGRQGLFVAKKK 473
Db 442 LRKLETVKEKDVIFISDFSEEDYNDIVGMDKLRATKAGRQGLFVAKKK 494

RESULT 2
Q9AXH3 PRELIMINARY; PRT; 491 AA.
ID Q9AXH3
AC Q9AXH3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Phosphoethanolamine N-methyltransferase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Parani M., Parida A.;
RT "Characterization of a cDNA for phosphoethanolamine N-methyltransferase."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF228858; AGS9894.1; --
DR GO: GO:0005344; P:cytosol; F:protein transport activity; IEA.
DR GO: GO:0008757; P:S-adenosylmethionine-dependent methyltransferase; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0006810; P:transferase; IEA.
DR InterPro: IPR000896; Hemocyanin.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000051; SAM bind.
DR PROSITE: PS00210; HEMOCYANIN.2; 1.
DR Methyltransferase, Transferase.
SQ SEQUENCE 491 AA; 55947 MW; D36RCDBS12733E7 CRC64;
Query Match 5.7%; Score 27; DB 10; Length 491;
Best Local Similarity 100.0%; Pred. No. 4.6e-19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 FVIVSRDTILHIQDKPALFRSFKYKL 357
Db 349 FVIVSRDTILHIQDKPALFRSFKYKL 375

RESULT 3
Q8L7A8 PRELIMINARY; PRT; 376 AA.
ID Q8L7A8
AC Q8L7A8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Phosphoethanolamine N-methyltransferase, putative.
GN ATIG73600.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosida II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick P., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Cheuk H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Shinn P., Tang C.C., Toroumi M., Wallender B.K., Wong C.,
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J.,
RA Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY136372; AA97038.1; --
DR GO: GO:0008757; P:S-adenosylmethionine-dependent methyltransferase; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000051; SAM bind.
KW Transferase; Methyltransferase.
SQ SEQUENCE 376 AA; 42811 MW; 2B4F6BE2B112B3ED CRC64;

Query Match 5.5%; Score 26; DB 10; Length 376;
Best Local Similarity 100.0%; Pred. No. 4.2e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 MLDLKPQKQVLDVGGGIGGDFYMAE 284
Db 290 MLDLKPQKQVLDVGGGIGGDFYMAE 315

RESULT 4
Q84SA4 PRELIMINARY; PRT; 493 AA.
ID Q84SA4
AC Q84SA4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphoethanolamine N-methyltransferase.
OS Aster tripolium (Sea aster).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Campanulids; Asterales; Asteraceae; Asteroideae; Asteraceae; Aster.
OX NCBI_TaxID=74787;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeda M., Uno Y., Kanachi M., Inagaki N.;
RT "Analysis of nine cDNAs for salt-inducible genes in the halophyte sea aster."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB090883; BACS7960.1; --
DR GO: GO:0008757; P:S-adenosylmethionine-dependent methyltransferase; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000051; SAM bind.
DR Methyltransferase, Transferase.
SQ SEQUENCE 493 AA; 56002 MW; 9F2C7369192B6DA5 CRC64;

Query Match 5.5%; Score 26; DB 10; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 YPDNSFDVYSRDTILHIQDKPALFR 351
Db 346 YPDNSFDVYSRDTILHIQDKPALFR 371

RESULT 5
Q7XJ2

ID Q7XJ2 PRELIMINARY; PRT; 491 AA.
 AC Q7XJ2, 01-JUN-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Phosphoethanolamine N-methyltransferase.
 OS Brassica napus (rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ye C., Li J., Yang J., Wang B.;
 RT "Plant gene cloning and expression analysis under different
 RL stresses";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY319479; AAP3582.1; -;
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR00051; SAM bind.
 SQ SEQUENCE 491 AA; 55938 MW; A08C3318737031EE CRC64;
 Query Match 5.1%; Score 24; DB 10; Length 491;
 Best Local Similarity 100.0%; Pred. No. 6.7e-16; Mismatches 0; Indels 0; Gaps 0;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 261 DLKPGQKVLVGGCGGDFYMAE 284
 DB 279 DLKPGQKVLVGGCGGDFYMAE 302
 RESULT 6
 Q8VX1 PRELIMINARY; PRT; 498 AA.
 AC Q8VX1, 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Phosphoethanolamine methyltransferase.
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Frenette Charron J.-B., Breton G., Danyluk J., Muzac I., Ibrahim R.,
 RA Sarhan F.;
 RT "Molecular and biochemical characterization of a cold regulated
 RT phosphoethanolamine methyltransferase from wheat";
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY065971; AAL40895.1; -;
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR00051; SAM bind.
 KW Transferase; Methyltransferase.
 SQ SEQUENCE 498 AA; 56858 MW; 37BB7134E2DEA148 CRC64;
 Query Match 4.4%; Score 21; DB 10; Length 498;
 Best Local Similarity 100.0%; Pred. No. 9.5e-13; Mismatches 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 265 GQKVLVGGCGGDFYMAE 285
 DB 291 GQKVLVGGCGGDFYMAE 311
 RESULT 7
 Q8LJ10 PRELIMINARY; PRT; 499 AA.
 AC Q8LJ10, 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Putative phosphoethanolamine methyltransferase.
 GN P0431H09.20.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0431H09.20";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003248; BAC10708.1; -;
 DR Gramineae; Oryzoideae; Oryza.
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR00051; SAM bind.
 KW Transferase; Methyltransferase.
 SQ SEQUENCE 499 AA; 56786 MW; ADDF3D04E62D18BE CRC64;
 Query Match 3.6%; Score 17; DB 10; Length 499;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08; Mismatches 0; Indels 0; Gaps 0;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 331 FDVIYSRDTLHIOQKP 347
 DB 358 FDVIYSRDTLHIOQKP 374
 RESULT 8
 Q82720 PRELIMINARY; PRT; 346 AA.
 AC Q82720, 01-NOV-1998 (TRENBLrel. 08, Created)
 DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE S-adenosyl-methionine-sterol-C-methyltransferase.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=xanthi SH6;
 RX MEDLINE=98417432; PubMed=9746350;
 RA Bouvier-Nave P., Husselstein T., Benveniste P.;
 RT "Two families of sterol methyltransferases are involved in the first
 RT and the second methylation steps of plant sterol biosynthesis";
 RL Eur. J. Biochem. 256:189-96 (1998).
 DR EMBL; U01312; AAC34951.1; -;
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR00051; SAM bind.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 346 AA; 38871 MW; E7A3F9534A9D94A4 CRC64;
 Query Match 3.4%; Score 16; DB 10; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07; Mismatches 0; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 262 LKPGQKVLVGGCGIGG 277
 DB 98 LKPGQKVLVGGCGIGG 113
 RESULT 9
 ID Q82434 PRELIMINARY; PRT; 349 AA.

AC 082434;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE S-adenosyl-methionine cycloartenol-C24-methyltransferase.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=xanthi SH6; TISSUE=Callus;
RX MEDLINE=98417432; PubMed=9746350;
RA Bouvier-Nave P., Huselestein T., Benveniste P.;
RT "Two families of sterol methyl-transferases are involved in the first
and the second methylation steps of plant sterol biosynthesis.";
RL Eur. J. Biochem. 256:88-96(1998).
RL EMBL; AF053766; AAC35787.1; -
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000051; SAM bind.
KW Methyltransferase; Transferase.
SQ SEQUENCE 349 AA; 39028 MW; ED6A5853CABDCD77 CRC64;

Query Match 3.4%; Score 16; DB 10; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.2e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 16; Conservative 0;

QY 262 LKPGQKVLVDVGGGIGG 277
DB 99 LKPGQKVLVDVGGGIGG 114

RESULT 10
Q43445 PRELIMINARY; PRT; 367 AA.
ID Q43445
AC Q43445
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE S-adenosyl-L-methionine:delta24-sterol-C-methyltransferase.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams 82; TISSUE=Etisolated hypocotyl;
RX MEDLINE=96199190; PubMed=8621604;
RA Shi J., Gonzales R.A., Bhattacharya M.K.;
RT "Identification and characterization of an S-adenosyl-L-methionine:
delta 24-sterol-C-methyltransferase cDNA from soybean.";
RL J. Biol. Chem. 271:9384-9389(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams 82; TISSUE=Etisolated hypocotyl;
RA Clouse J.A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U43683; AAB04057.1; -
DR PIR; T06780; T06780.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000051; SAM bind.
KW Methyltransferase; Transferase.
SQ SEQUENCE 367 AA; 41518 MW; B3BEP8D2F0B22FB4 CRC64;

Query Match 3.4%; Score 16; DB 10; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.2e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 16; Conservative 0;

QY 262 LKPGQKVLVDVGGGIGG 277

DB 119 LKPGQKVLVDVGGGIGG 134

RESULT 11
Q8LKM1 PRELIMINARY; PRT; 336 AA.
ID Q8LKM1
AC Q8LKM1
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Cephalopod.
GN CPH.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12100483;
RA Schrick K., Mayer U., Martin G., Bellini C., Kuhn C., Schmidt J.,
RA Jurgens G.;
RT "Interactions between sterol biosynthesis genes in embryonic
development of Arabidopsis.";
RL Plant J. 31:61-73(2002).
DR EMBL; AF494289; AAMS3553.1; -
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.
DR InterPro; IPR000051; SAM bind.
SQ SEQUENCE 336 AA; 38267 MW; BF52441B439B174 CRC64;

Query Match 3.0%; Score 14; DB 10; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.4e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

QY 264 PGQKVLVDVGGGIGG 277
DB 94 PGQKVLVDVGGGIGG 107

RESULT 12
Q9LM02 PRELIMINARY; PRT; 336 AA.
ID Q9LM02
AC Q9LM02
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE SAM:cycloartenol-C24-methyltransferase (24-sterol C-methyltransferase)
(Sterol methyltransferase SMT1).
GN SMT1 OR ATSG13710.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Schaeffer A., Schaller H., Benveniste P.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98069011; PubMed=9405937;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
Sequence features of the regions of 1,044,062 bp covered by thirteen
physically assigned P1 clones.";
RL DNA Res. 4:291-300(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

RA Diener A.C., Li H., Zhou W.-X., Whoriskey W.J., Nes W.D., Fink G.R.;
 RT "Effects of reduced C-24 sterol alkyltion on plant growth due to a
 RT deficiency in sterol methyltransferase 1";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 (4)
 RN SEQUENCE FROM N.A.
 RP Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Naruaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 (5)
 RN SEQUENCE FROM N.A.
 RP Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Naruaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF090372; AAC78847.1; -;
 DR EMBL: AB006704; BAB08698.1; -;
 DR EMBL: AF195648; AAC28462.1; -;
 DR EMBL: AY120716; AAM53374.1; -;
 DR EMBL: BT000058; AAM15377.1; -;
 DR GO: GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
 DR GO: GO:0016740; P:transferase activity; IEA.
 DR InterPro: IPR000051; SAM bind.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 336 AA; 38268 MW; 46498B368D81CE9 CRC64;

Query Match 3.0%; Score 14; DB 10; Length 336;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 PQQVLDVGGGIGG 277
 DB 94 PQQVLDVGGGIGG 107

RESULT 13
 O24328 PRELIMINARY; PRT; 346 AA.
 ID O24328
 AC O24328
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE S-adenosyl-methionine-sterol-C-methyltransferase.
 OS Ricinus communis (Castor bean)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosoids I; Malpighiales; Euphorbiaceae; Acalyphaceae; Acalyphaceae;
 OC Ricinus.
 OX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kaker 296;
 RX MEDLINE=97352551; PubMed=9708946;
 RA Bouvier-Nave P., Huselestein T., Desprez T., Benveniste P.;
 RT "Identification of cDNAs encoding sterol methyl-transferases involved
 RT in the second methylation step of plant sterol biosynthesis";
 RL Eur. J. Biochem. 246:518-529 (1997).
 DR EMBL: U01313; AAB62812.1; -;
 DR EMBL: T10173; T10173.
 DR GO: GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
 DR GO: GO:0016740; P:transferase activity; IEA.
 DR InterPro: IPR000051; SAM bind.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 346 AA; 38930 MW; 96F3DA79CA81AE82 CRC64;

Query Match 2.5%; Score 12; DB 10; Length 346;

Best Local Similarity 100.0%; Pred. No. 0.0018;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 266 QKVLVDVGGGIGG 277
 DB 102 QKVLVDVGGGIGG 113

RESULT 14
 Q84M50 PRELIMINARY; PRT; 330 AA.
 ID Q84M50
 AC Q84M50
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative endosperm C-24 sterol methyltransferase.
 GN OSJNBA0059E14.10.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ccv. Nipponbare;
 RA Buell C.R., Yuan Q., Cuiyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Fadrosah D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNBA0059E14 genomic sequence";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 (2)
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ccv. Nipponbare;
 RA Buell R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC135958; AAP21419.1; -;
 DR GO: GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
 DR GO: GO:0016740; P:transferase activity; IEA.
 DR InterPro: IPR000051; SAM bind.
 KW Methyltransferase; Transferase.
 SQ SEQUENCE 330 AA; 36900 MW; 1ADC1039C3E73C7D CRC64;

Query Match 2.3%; Score 11; DB 10; Length 330;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLVDVGGGIGG 277
 DB 85 KVLVDVGGGIGG 95

RESULT 15
 P93852 PRELIMINARY; PRT; 344 AA.
 ID P93852
 AC P93852
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Endosperm C-24 sterol methyltransferase.
 GN ESMT1
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97435974; PubMed=9290641;
 RA Grebenok R.J., Galbraith D.W., Penna D.D.;
 RT "Characterization of Zea mays endosperm C-24 sterol methyltransferase";

RT one of two types of sterol methyltransferase in higher plants.";
RL Plant Mol. Biol. 34:891-896(1997).
DR EMBL; U79669; AAB70886.1; -.
DR PIR; T04138; T04138.
DR GO; GO:0008757; P:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR000051; SM bind.
KW Methyltransferase; Transferase.
SQ SEQUENCE 344 AA; 38779 MW; 2794A4585B29EAB CRC64;

Query Match 2.3%; Score 11; DB 10; Length 344;
Best Local Similarity 100.0%; Pred. NO. 0.021;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 KVLGVGGGIGG 277
|||
Db 103 KVLGVGGGIGG 113

Search completed: July 26, 2004, 13:23:30
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:10:26 ; Search time 54 seconds
(without alignments)
2474.907 Million cell updates/sec

Title: US-10-031-331B-40
Perfect score: 2515
Sequence: 1 HTVDLTIEMMLDSQASDLID.....KLRRTAGEQWGLFVAKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_29Jan04:*

1: Geneseq1980e:*

2: Geneseq1990a:*

3: Geneseq2000b:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2515	100.0	473	4	AAB80627
2	2204	87.6	494	4	AAB09760
3	1829.5	72.7	491	4	AAB99889
4	1270	50.5	289	4	AAB09761
5	1033	41.1	287	7	ABM74462
6	779	31.0	191	7	ABM74064
7	658	26.2	168	4	AAB99873
8	241.5	9.6	281	6	AAB35493
9	225.5	9.0	276	4	ABG99864
10	224	8.9	280	6	ABP96395
11	224	8.9	280	7	ADD19156
12	218.5	8.7	283	3	AAB32507
13	218.5	8.7	283	7	ADE10279
14	214	8.5	280	6	ABP96396
15	214	8.5	280	7	ADD19157
16	212.5	8.4	346	5	AB13618
17	210.5	8.4	317	2	AAM95016
18	210.5	8.4	317	6	ABP96397
19	209.5	8.3	285	3	AAY70038
20	200.5	8.0	279	3	AAY70041
21	200.5	8.0	345	6	ABP96378
22	200.5	8.0	345	7	ADD19145
23	199.5	7.9	283	6	ABP57689
24	198	7.9	344	3	AAT60392
25	194.5	7.7	283	2	AAY39308

26	194.5	7.7	299	4	AAB70954
27	191	7.6	299	6	ABP96392
28	190	7.6	299	6	ABM6984
29	189.5	7.5	376	6	ABP96379
30	189.5	7.5	376	7	ADD19146
31	187.5	7.5	330	6	ABP96399
32	187.5	7.5	336	3	AAQ21651
33	187.5	7.5	339	3	AAQ21525
34	187.5	7.5	361	3	AAQ21524
35	187.5	7.5	366	3	AAQ21650
36	187.5	7.5	370	4	AAB99888
37	187	7.4	293	3	AAQ21652
38	187	7.4	296	3	AAQ21526
39	186.5	7.4	314	5	ABP93454
40	186.5	7.4	315	6	ABP96391
41	186.5	7.4	336	3	AAQ45951
42	186.5	7.4	336	3	AAQ45954
43	186.5	7.4	358	3	AAQ45950
44	186.5	7.4	365	3	AAQ45953
45	186	7.4	293	3	AAQ45955

ALIGNMENTS

RESULT 1

AAB80627
ID AAB80627 standard; protein; 473 AA.

XX AAB80627;

DT 06-AUG-2003 (revised)

DT 02-MAY-2001 (first entry)

XX Environmental stress tolerant protein SEQ ID 40.

XX Environmental stress resistance; salt; heat; desert; transgenic plant.

OS Suaeda japonica.

PN WO200106006-A1.

XX 25-JAN-2001.

PF 19-JUL-2000; 2000WO-JP004862.

PR 19-JUL-1999; 99JP-00235910.

PR 24-MAR-2000; 2000JP-00085377.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Yamada A, Ozeki Y, Saito T;

XX WPI: 2001-147355/15.

XX N-PSDB; AAF74206.

XX Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.

XX Claim 64; Page 125-127; 167pp; Japanese.

XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining DNA encoding environmental stress resistance factors. The DNA encoding proteins conferring environmental stress resistance can be used in the production of plants resistant to environmental stress, which can be cultivated in unfavourable

CC environments such as deserts, salt damaged ground, cold regions and the
 CC oceans. They can be used for increasing the area of land covered by green
 CC plants, and desert greening and afforestation, in order to counter the
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR
 CC primers AAF74219 and AAF74220 are used in an example illustrating the
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 473 AA;

Query Match 100.0%; Score 2515; DB 4; Length 473;
 Best Local Similarity 100.0%; Pred. No. 3.9e-237;
 Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HTVDLTIANMLDSQASDLKERRPEITLSMLPLEGKCLLELGAGIGRTGLAQAQV 60
 DB 1 HTVDLTIANMLDSQASDLKERRPEITLSMLPLEGKCLLELGAGIGRTGLAQAQV 60

QY 61 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLPPPHSLDVIKSNMMLYLSDEVE 120
 DB 61 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLPPPHSLDVIKSNMMLYLSDEVE 120

QY 121 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 180
 DB 121 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 180

QY 181 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 240
 DB 181 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 240

QY 241 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 300
 DB 241 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 300

QY 301 SPALERSIGLKCAVEFEVADCTKINYPNSFDVIYSRDTTLIHQDKPALFRSFYKWLKPG 360
 DB 301 SPALERSIGLKCAVEFEVADCTKINYPNSFDVIYSRDTTLIHQDKPALFRSFYKWLKPG 360

QY 361 GKVLSDYCKKAGPPSPFAAYIKORGVDLHDVKYQOMLKDAGFDVLAERTQPIRV 420
 DB 361 GKVLSDYCKKAGPPSPFAAYIKORGVDLHDVKYQOMLKDAGFDVLAERTQPIRV 420

QY 421 LRKELETVEKEDVFISDFSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKK 473
 DB 421 LRKELETVEKEDVFISDFSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKK 473

RESULT 2

AAE09760
 ID AAE09760 standard; protein; 494 AA.

XX
 AC AAE09760;

XX
 DT 29-NOV-2001 (first entry)

XX
 DE Spinach PEAMT protein.

XX Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
 KW glycine betaine; choline-O-sulphate; lipid content alteration;
 KW osmotic stress tolerance; nutritional value; transgenic plant;
 KW cryoprotectant.

XX Spinacia oleracea.

XX Key Location/Qualifiers

XX Key

XX Misc-difference 462

XX Note="This residue is given as Lys in the sequence
 shown as SEQ ID NO: 2 in figure 3 of the specification"

XX W0200168870-A2.

XX W0200168870-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-US008352.
 XX
 PR 15-MAR-2000; 2000US-00525885.
 XX
 PA (UYFL) UNIV FLORIDA.
 PA (UYCA-) UNIV CARNEGIE MELLON.
 XX
 PI Hanson AD, Nuccio ML, Henry SA;
 XX
 DR WPI; 2001-565796/63.
 DR N-PSDB; AAD16797.
 XX
 PT New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
 PT polypeptides, useful for modulating the levels of cellular intermediates
 PT such as phosphodimethylethanolamine and for altering the lipid content in
 PT plants cells.
 XX
 XX Claim 1; Page 109; 158pp; English.
 XX
 XX The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-
 CC methyltransferase (PEAMT) protein from spinach. The PEAMT sequences are
 CC useful for modulating the levels of cellular intermediates such as
 CC phosphodimethylethanolamine, phosphomono-methylethanolamine, choline,
 CC phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine
 CC betaine. They are useful for altering the lipid content in plant cells.
 CC The polynucleotides are also useful for improving the osmotic stress
 CC tolerance of a plant and increasing the cryoprotectant properties of a
 CC plant. The present invention also relates to methods and compositions
 CC comprising PEAMT used for generating transgenic plants with increased
 CC nutritional value
 XX
 SQ Sequence 494 AA;

Query Match 87.6%; Score 2204; DB 4; Length 494;
 Best Local Similarity 86.4%; Pred. No. 1.3e-206;
 Matches 408; Conservative 34; Mismatches 30; Indels 0; Gaps 0;

QY 1 HTVDLTIANMLDSQASDLKERRPEITLSMLPLEGKCLLELGAGIGRTGLAQAQV 60
 DB 22 HSDVLTIANMLDSQASDLKERRPEITLSMLPLEGKCLLELGAGIGRTGLAQAQV 81

QY 61 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLPPPHSLDVIKSNMMLYLSDEVE 120
 DB 82 IALDFTESAIKKNVEINGHYKNVKNPCADVTSTLPPPHSLDVIKSNMMLYLSDEVE 141

QY 121 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 180
 DB 142 NLVERMLKWLKPGGYIPFRESCTHQSGDHKRNKNTHTYRPPRYTKAFKCHLQDGSNS 201

QY 181 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 240
 DB 202 YELSLSCCKTCGAVVRNKNQNOISMLWQVDSKDKGFORFLDTSTQYKCNILRYERVF 261

QY 241 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 300
 DB 262 GPGYVSTGGYETTKFVSMILDKPGQKVLVGGGIGGDFYMAETFDVWVGFDSLVSNNI 321

QY 301 SPALERSIGLKCAVEFEVADCTKINYPNSFDVIYSRDTTLIHQDKPALFRSFYKWLKPG 360
 DB 322 SPALERSIGLKCAVEFEVADCTKINYPNSFDVIYSRDTTLIHQDKPALFRSFYKWLKPG 381

QY 361 GKVLSDYCKKAGPPSPFAAYIKORGVDLHDVKYQOMLKDAGFDVLAERTQPIRV 420
 DB 382 GKVLSDYCKKAGPPSPFAAYIKORGVDLHDVKYQOMLKDAGFDVLAERTQPIRV 441

QY 421 LRKELETVEKEDVFISDFSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKK 472
 DB 442 LRKELETVEKEDVFISDFSEEDYNDIVGGNDKLRRTAKGEQRMGLFVAKK 493

RESULT 3
 AAB99889

ID AAB99889 standard; protein; 491 AA.
 XX AAB99889;
 AC 21-SEP-2001 (first entry)
 DT
 DT
 DT
 DE Physcomitrella patens 78_pprot1_092_e12-260rev protein.
 XX
 XX Tocopherol and carotenoid metabolism related protein; TCMP; synthesis;
 KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;
 KW identification; genome mapping; modulation; evolutionary study;
 KW cellular production; fine chemical.
 XX
 XX Physcomitrella patens.
 XX WO200144276-A2.
 XX
 XX 21-JUN-2001.
 XX
 XX 14-DEC-2000; 2000WO-EP012698.
 XX
 XX 16-DEC-1999; 99US-0171121P.
 XX
 XX (BADI) BASF PLANT SCI GMBH.
 XX
 XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;
 PI
 PI
 PI
 DR WPI; 2001-398121/42.
 DR
 DR
 XX Tocopherol and carotenoid metabolism related protein (TCMP), used to
 PT produce fine chemicals, is isolated from mosses, algae, microorganisms,
 PT fungi, plants, or their fragments.
 XX
 XX Claim 28; Page 122-123; 123pp; English.
 XX
 XX The present invention describes isolated tocopherol and carotenoid
 CC metabolism related proteins (TCMP) (I) from mosses or algae,
 CC microorganisms or fungi, plants, or its fragments. (I) can be used as
 CC enzymes in the production of fine chemicals or in the metabolism of
 CC tocopherols and carotenoids. (I) also assist in transmembrane transport.
 CC The fine chemicals that can be produced include lipids, fatty acids,
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.
 CC Nucleotide sequences, proteins, vectors and host cells from the present
 CC invention can be used: (a) to identify mosses related to Physcomitrella
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella patens;
 CC (c) in the modulation of functional TCMP regions; (d) in evolutionary studies;
 CC (e) in the determination of functional TCMP regions; (f) and in the
 CC cellular production of fine chemicals. AAH44222 to AAH44262 encode the
 CC Physcomitrella patens TCMP proteins given in AAB99849 to AAB99889.
 CC AAH44212 to AAH44221 represent nucleotide sequence used in the
 CC exemplification of the present invention
 XX
 XX Sequence 491 AA;
 SQ
 Query Match 72.7%; Score 1829.5; DB 4; Length 491;
 Best Local Similarity 70.0%; Pred. No. 6.4e-170;
 Matches 332; Conservative 65; Mismatches 74; Indels 3; Gaps 2;
 OY 1 HTVLTIRAWLDSQASOLDKEEPILSLMLPPLGKCLLELGANGRTFGLAEKAGOV 60
 DB 17 HSVEPSVAMLDQSKSLDKERPEILSLPPLPPKNDKWLANGRTFGLAKUAGHV 76
 OY 61 IALDPTESATKNEVINGXNVKPMCAVDTSPILSPFPHSLDVFNSMLMLTDEEVE 120
 DB 77 LAMDFWENLIKKNEDVNGHNNIDFKCADVTSPLDNTAAGSADIVFSNMLTLDSEYK 136
 OY 121 NLVEMLMKLPKGGYLFPPFSCFHSGDHKKSNPTHYRPFYKAKPKCHLDQSGNS 180
 DB 137 GLASRWENLRPGYLFPPFSCFHSGDHKKSNPTHYRQENYTNFQAYIEE-DGSY 195
 OY 181 YELSLSCCKIGAVNKKKNQNOISWLMQV--DSKDDKGFQFLDTSQYKCNLSILYER 238

Db 196 FREWVGCCKCVTYVRNKNQVCMWRKYSQSDGPESECFKFLDQOYTSTGILRYER 255
 OY 239 VFGPGVSTGYETTKBFVSMLDLKKGOKVLDVCGGIGGGDFYMASTFVDEVVGFDSVN 298
 Db 256 IFGGFVSTGGIETTRAFVSMLDLKKGOKVLDVCGGIGGGDFYMASTFVDEVVGFDSVN 315
 OY 299 MISFALERSIGLKCAVEFEVADCTKINTPNSPVIYSTRDILHIQKPKALFSPYKWLK 358
 Db 316 MISFALERSIGRKCAVEFEVGDCTKINYPHASFVIVSTRDILHIQKPKALFSPYKWLK 375
 OY 359 PGCKVLISDYCKKAGPPSPFEFAAYIKQGYLDHVDVKEYGQMLKADAGFVDVLAEDRTEQFI 418
 Db 376 PGGRVLISDYCRAPQTPSAEFAAYIQQGYDLNHSVQKYGEMLEADAGFVFEVVAEDRTEQFI 435
 OY 419 RVLRKELETVEKQVFISSDEEDYNDIVGWNDKLRRRTAKGRWGLFVANK 472
 Db 436 EVLQRELATTEAGRDQFINDFSEEDYNIYVGWNSKLRCSNDEQKGLFIANK 489
 RESULT 4
 AAE09761
 ID AAE09761 standard; protein; 289 AA.
 XX
 AC AAE09761;
 DT 29-NOV-2001 (first entry)
 XX
 DE Spinach PEAMT truncated protein.
 XX
 KW Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase;
 KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;
 KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
 KW glycine betaine; choline-O-sulphate; lipid content alteration;
 KW osmotic stress tolerance; nutritional value; transgenic plant;
 KW cryoprotectant.
 XX
 OS Spinacia oleracea.
 XX WO200168870-A2.
 XX
 XX 20-SEP-2001.
 XX
 XX 15-MAR-2001; 2001WO-US008352.
 XX
 XX 15-MAR-2000; 2000US-00525885.
 XX
 XX (UYFL) UNIV FLORIDA.
 XX (UYCA-) UNIV CARNEGIE MELLON.
 XX
 XX Hanson AD, Nuccio ML, Henry SA;
 XX WPI; 2001-565796/63.
 XX N-PSDB; AAD16798.
 XX
 XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
 PT polypeptides, useful for modulating the levels of cellular intermediates
 PT such as phosphodimethylethanolamine and for altering the lipid content in
 PT plants cells.
 XX
 XX Claim 1; Page 109; 158pp; English.
 PS
 XX The present sequence is S-adenosyl-L-methionine:phosphoethanolamine N-
 CC methyltransferase (PEAMT) truncated protein from spinach. The PEAMT
 CC sequences are useful for modulating the levels of cellular intermediates
 CC such as phosphodimethylethanolamine, phosphomono-methylethanolamine,
 CC choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or
 CC glycine betaine. They are useful for altering the lipid content in plant
 CC cells. The polynucleotides are also useful for improving the osmotic
 CC stress tolerance of a plant and increasing the cryoprotectant properties
 CC of a plant. The present invention also relates to methods and
 CC compositions comprising PEAMT used for generating transgenic plants with
 CC increased nutritional value

XX SQ Sequence 289 AA;
 Query Match 50.5%; Score 1270; DB 4; Length 289;
 Best Local Similarity 88.7%; Pred. No. 2e-115;
 Matches 235; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

QY 1 HTVDLTIAAMLDQSDLDKPERPEILSMPLPEKCKLLELGAGIGRTGELAKAGQV 60
 DB 22 HSDVLTVEAMMLDQSDLDKPERPEILSMPLPEKCKLLELGAGIGRTGELAKAGQV 81
 QY 61 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFPHSLDVFPSNMLLYLSDEVE 120
 DB 82 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFPHSLDVFPSNMLLYLSDEVE 141
 QY 121 NLVERMLKWLKPGGYIPFRESCHQSDHKKKNPHTHYREPRFYTKAFKCHLODGSNS 180
 DB 142 NLVERMLKWLKPGGYIPFRESCHQSDHKKKNPHTHYREPRFYTKAFKCHLODGSNS 201
 QY 181 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORDTDSQYKNSILYERVP 240
 DB 202 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORDTDSQYKNSILYERVP 261
 QY 241 GFGYVSTGGYETTKFVSMLDLKP 265
 DB 262 GFGYVSTGGYETTKFVSMLDLKP 286

RESULT 5
 ABM74462
 ID ABM74462 standard; protein; 287 AA.
 AC ABM74462;
 XX 17-OCT-2003 (first entry)
 XX DNA clone originating in barley containing SNP sequence #872.
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
 XX Hordeum vulgare.
 XX WO2003057877-A1.
 XX 17-JUL-2003.
 XX 16-DEC-2002; 2002WO-IB005403.
 XX 20-DEC-2001; 2001JP-00387059.
 XX 20-DEC-2001; 2001JP-00387131.
 XX 20-DEC-2001; 2001JP-00403299.
 XX 20-DEC-2001; 2001JP-00403300.
 XX 27-SEP-2002; 2002JP-00327515.
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 XX Sato K, Takeda K, Kohara Y;
 XX WPI; 2003-587127/55.
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 XX sequences containing them for analysis and identification of barley
 XX varieties and production of barley transformants with desired
 XX characteristics.
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 XX The present invention relates to oligonucleotide clones originating in
 XX barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 XX (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 XX varieties, identification of particular varieties and genotype-phenotype
 XX analysis, isolation of specific genes and creation of new varieties by
 XX transformation of barley varieties with them and production of new barley

XX varieties with desired properties. The present sequence represents an
 XX oligonucleotide clone sequence featured in the specification. The
 XX sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published-pct-sequences
 XX SQ Sequence 287 AA;
 Query Match 41.1%; Score 1033; DB 7; Length 287;
 Best Local Similarity 76.9%; Pred. No. 3.3e-92;
 Matches 193; Conservative 24; Mismatches 32; Indels 0; Gaps 0;

QY 1 HTVDLTIAAMLDQSDLDKPERPEILSMPLPEKCKLLELGAGIGRTGELAKAGQV 60
 DB 37 HSDVLTVEAMMLDQSDLDKPERPEILSMPLPEKCKLLELGAGIGRTGELAKAGV 96
 QY 61 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFPHSLDVFPSNMLLYLSDEVE 120
 DB 97 IALDFIRSAIKKNEVINGHYKVKFMCADVTSPFPHSLDVFPSNMLLYLSDEVE 156
 QY 121 NLVERMLKWLKPGGYIPFRESCHQSDHKKKNPHTHYREPRFYTKAFKCHLODGSNS 180
 DB 157 NLVERMLKWLKPGGYIPFRESCHQSDHKKKNPHTHYREPRFYTKAFKCHLODGSNS 216
 QY 181 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORDTDSQYKNSILYERVP 240
 DB 217 YELSLSCCKCIGAVKKNQNOISWLWQVDSKDDKGFORDTDSQYKNSILYERVP 276
 QY 241 GFGYVSTGGYETTKFVSMLDLKP 251
 DB 277 GFGYVSTGGYETTKFVSMLDLKP 287

RESULT 6
 ABM74064
 ID ABM74064 standard; protein; 191 AA.
 AC ABM74064;
 XX 17-OCT-2003 (first entry)
 XX DNA clone originating in barley containing SNP sequence #474.
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.
 XX Hordeum vulgare.
 XX WO2003057877-A1.
 XX 17-JUL-2003.
 XX 16-DEC-2002; 2002WO-IB005403.
 XX 20-DEC-2001; 2001JP-00387059.
 XX 20-DEC-2001; 2001JP-00387131.
 XX 20-DEC-2001; 2001JP-00403299.
 XX 20-DEC-2001; 2001JP-00403300.
 XX 27-SEP-2002; 2002JP-00327515.
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 XX Sato K, Takeda K, Kohara Y;
 XX WPI; 2003-587127/55.
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 XX sequences containing them for analysis and identification of barley
 XX varieties and production of barley transformants with desired
 XX characteristics.
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 XX The present invention relates to oligonucleotide clones originating in

CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences
 XX
 XX Sequence 191 AA;

Query Match 31.08; Score 779; DB 7; Length 191;
 Best Local Similarity 73.8%; Pred. NO. 1.4e-67;
 Matches 141; Conservative 26; Mismatches 24; Indels 0; Gaps 0;
 QY 282 MAETFDVEVGFDSVNMISFALERSIGLKCAVEFEVADCTKINYPDINSFDVIYSRDTIL 341
 DB 1 MAENYDVHVVGIDLSINNVSFALHAIGRKAIVEFEVADCTTKYPTDNTFDVIYSRDTIL 60
 QY 342 HQDKPALFRSPYKWLPGGKVLISDYCKKAGPSPPEFAAYTKQRYDLHDVKEYGWLK 401
 DB 61 HQDKPALFRSPYKWLPGGKVLISDYCKSPGPEEFASYTKQRYDLHDVKEYGWLK 120
 QY 402 DAGFVDVLAEDRTSQTRIVRLKSLTVEKEKDVFSDFSEEDYNDIVGWNKDLRTAKG 461
 DB 121 NAGFHDVIAEDRSQFLKVLQRLAEVKNKDFLADFGQEDYDIDVTGMNAKLQSSAG 180
 QY 462 EQRWGLFVAKK 472
 DB 181 EQRWGLFVGTK 191

RESULT 7
 AAB99873
 ID AAB99873 standard; protein; 168 AA.
 AC AAB99873;
 DT 21-SEP-2001 (first entry)
 DE Physcomitrella patens 7a_ppprotcl_092_e12rev protein.
 XX Tocopherol and carotenoid metabolism related protein; TCWRP; synthesis;
 KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;
 KW identification; genome mapping; modulation; evolutionary study;
 KW cellular production; fine chemical.
 XX Physcomitrella patens.
 OS
 PN W0200144276-A2.
 PD 21-JUN-2001.
 XX 14-DEC-2000; 2000WO-EP012698.
 XX 16-DEC-1999; 99US-0171121P.
 XX (BADI) BASF PLANT SCI GMBH.
 XX Lerchl J, Benz A, Ehrhardt T, Reindl A, Cirpus P, Bieschoff P;
 PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;
 XX WPI; 2001-398121/42.
 DR N-PSDB; AAB44246.
 XX Tocopherol and carotenoid metabolism related protein (TCWRP), used to
 PT produce fine chemicals, is isolated from mosses, algae, microorganisms,
 PT fungi, plants, or their fragments.
 XX Claim 28; Page 118-119; 123pp; English.

CC The present invention describes isolated tocopherol and carotenoid
 CC metabolism related proteins (TCWRP) (I) from mosses or algae,
 CC microorganisms or fungi, plants, or its fragments. (I) can be used as
 CC enzymes in the production of fine chemicals or in the metabolism of
 CC tocopherols and carotenoids. (I) also assist in transmembrane transport.
 CC The fine chemicals that can be produced include lipids, fatty acids,
 CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.
 CC Nucleotide sequences, proteins, vectors and host cells from the present
 CC invention can be used: (a) to identify mosses related to Physcomitrella
 CC patens; (b) in mapping genomes of mosses related to Physcomitrella patens;
 CC (c) in the modulation of TCWRP activity; (d) in evolutionary studies;
 CC (e) in the determination of functional TCWRP regions; (f) and in the
 CC cellular production of fine chemicals. AAB44222 to AAB44262 encode the
 CC Physcomitrella patens TCWRP proteins given in AAB99849 to AAB99889.
 CC AAB44212 to AAB44221 represent nucleotide sequence used in the
 CC exemplification of the present invention
 XX
 XX Sequence 168 AA;

Query Match 26.2%; Score 658; DB 4; Length 168;
 Best Local Similarity 72.9%; Pred. NO. 8.1e-56;
 Matches 121; Conservative 19; Mismatches 26; Indels 0; Gaps 0;
 QY 307 SIGLKCAVEFEVADCTKINYPDINSFDVIYSRDTILHIQDKPALFRSPYKWLPGGKVLIS 366
 DB 1 SIGLKCAVEFEVADCTKINYPDINSFDVIYSRDTILHIQDKPALFRSPYKWLPGGKVLIS 60
 QY 367 DYCKKAGPSPPEFAAYTKQRYDLHDVKEYGWLKADGFVDVLAEDRTSQTRIVRLK 426
 DB 61 DYCRAPQTPSAEFAAYTKQRYDLHDVKEYGWLKADGFVDVLAEDRTSQTRIVRLK 120
 QY 427 TVEKEKDVFSDFSEEDYNDIVGWNKDLRTAKGQRWGLFVAKK 472
 DB 121 TTEAGRODFNDFSEEDYNYIVSGWKSCLKRCSNDEQKWLFIAYK 166

RESULT 8
 AAB35493
 ID AAB35493 standard; protein; 281 AA.
 AC AAB35493;
 DT 17-JUN-2003 (first entry)
 DE Streptomyces platensis subspecies rosaceus dorrigocin ORF10 protein.
 XX Polyketide biosynthesis; dorrigocin; DORR; lactimidomycin, LACT.
 KW Streptomyces platensis.
 OS
 PN W0200288176-A2.
 PD 07-NOV-2002.
 XX 26-APR-2002; 2002WO-CA000591.
 XX 26-APR-2001; 2001US-0286346P.
 XX (ECOP-) ECOPIA BIOSCIENCES INC.
 XX Farnet CM, Zazopoulos E, Staffa A, Yang X;
 XX WPI; 2003-201222/19.
 DR N-PSDB; AAD54217, AAD54227.
 XX Novel isolated or purified polypeptide involved in biosynthesis of
 PT polyketide dorrigocin or polyketide lactimidomycin, useful for preparing
 PT dorrigocin or lactimidomycin.
 XX Claim 13; Page 197-198; 312pp; English.
 XX The invention relates to novel proteins involved in the biosynthesis of
 CC polyketide dorrigocin (DORR) or lactimidomycin (LACT) biosynthesis by

microorganisms. Sequences of the invention allow direct manipulation of dorrigo, lactimidomycin and related chemical structures via chemical engineering of the enzymes involved in the biosynthesis of dorrigo and lactimidomycin. They are useful for introducing chemical handles into normally inert positions that permit subsequent chemical modifications and facilitate the development of polyketides. The genes and proteins of the invention can also be used to generate a focused library of analogues around a polyketide lead candidate to fine-tune the compound for optimal properties. They are useful for generating antibodies specific for the polyketide biosynthesis. The present sequence is *Streptomyces platensis* subspecies *roseaceus* DORR ORF10 protein

XX Sequence 281 AA;

Query Match 9.6%; Score 241.5; DB 6; Length 281;
Best Local Similarity 31.6%; Pred. No. 1.2e-14;
Matches 65; Conservative 40; Mismatches 80; Indels 21; Gaps 7;
QY 253 TKEFVSMDLKPGQKVLVGGGCGGDFYMAETFDVYVQPLSVNMI--SPALERSIGL 310
DB 63 TDTMTDLRIQCGQRVLVGGGCGGCPAMRTARTGAHVGTGAIKSDQIARATAGAGL 122
QY 311 KCAVEFEVADCTKINYPNSDFVIYSRDTILHIQDKPALFRSYKMLPGGKVLISDYCK 370
DB 123 SDREVEFHADAMELFPDDSDFAAIAIESIFHFMDRGRVLAIRRVLRGGRLVLTDPFE 182
QY 371 KAGPPSPFAAVIKQGYD--LHDV-----KEYGQMLKDG--FVDVLAEDRTEQPIR 419
DB 183 R-GFVPAE-----KQPAVDRLLRDPINTLARPEYVPMLEADGLRVELL--DITEQSVR 234
QY 420 VLKKELETVEKEKDVFTSDSEEDYN 445

DB 235 QTFFQMSQSGSQEMQTVFDDAEKFS 260

RESULT 9
ABG99864

ID ABG99864 standard; protein; 276 AA.

AC ABG99864;

XX 16-JAN-2003 (first entry)

DE S. cinnamomensis MonE/S-adenosylmethionine-dependent methyltransferase.

XX Monensin; gene cluster; polyketide synthase; antibiotic; antihelminthic;
XX insecticide; immunosuppressant; antifungal; antibacterial; polyether;
XX mon B; mon BII; mon CI; mon CII; mon H; mon I; mon J; mon K; mon L; mon M;
XX mon N.

XX Streptomyces cinnamomensis.

XX WO200168867-A1.

XX 20-SEP-2001.

XX 30-MAY-2000; 2000WO-GB002072.

XX 28-MAY-1999; 99GB-00012563.

XX (BIOT-) BIOTICA TECHNOLOGY LTD.

XX Leadlay PP, Staunton J, Olinyk M;

XX WPI; 2001-611393/70.

XX N-PSDB; ABX04971.

XX New DNA sequence encoding polyketide synthase, useful for the production
XX of polyketides such as antibiotic monensin.

XX Claim 6; Page 73; 21pp; English.

XX The invention relates to a DNA sequence which is a fully defined sequence

CC of 103551 base pairs appearing as ABX04971, or its variant, that it is
CC not a sequence encoding all or part amino acids 1-920 encoded by mon AI
CC as given in the specification. The DNA is the S. cinnamomensis polyketide
CC antibiotic monensin biosynthetic gene cluster. Also included are a
CC recombinant cloning or expression vector comprising the gene cluster, a
CC transformed host cell which has been transformed to contain the gene,
CC cluster (and is capable of expressing a corresponding polypeptide), a
CC hybridization probe derived from the gene cluster (for identification and
CC isolation of the same or analogous gene cluster, e.g. one which binds
CC specifically to a region of the monensin gene cluster selected from mon
CC BI, mon BII, mon CI, mon CII, mon H, mon I, mon J, mon K, mon L, mon M,
CC mon N), the use of the mon RI gene or variant and a monensin promoter to
CC control expression of a heterologous gene in *Streptomyces cinnamomensis*,
CC a polypeptide encoded by a portion of the monensin gene cluster
CC (preferably comprising mon BI, mon BII, mon AIX or mon AX or their
CC mutants, alleles or variants), an epoxide enzyme encoded by mon CI, a
CC cyclase enzyme encoded by mon CII, producing S. cinnamomensis capable of
CC enhanced levels of production of monensin comprising engineering it to
CC overexpress the mon RI gene, S. cinnamomensis containing multiple copies
CC of the mon RI gene and/or its variants, expressing a gene heterologous to
CC S. cinnamomensis comprising transforming S. cinnamomensis with DNA
CC encoding a heterologous gene and expressing the gene under control of the
CC activator gene mon RI or actII/orf4 and 13-propyl erythromycin A. The
CC processes and materials (enzyme systems, nucleic acids and vectors) are
CC useful for preparing polyketides by recombinant synthesis. The
CC polyketides are useful as insecticides, antibiotics, antihelminthics,
CC antifungals, antibacterials or other pharmaceuticals. In particular the
CC gene is useful for the production of monensin, an antibiotic polyether
CC polyketide. The present sequence represents a protein encoded by the
CC monensin gene cluster

XX Sequence 276 AA;

Query Match 9.0%; Score 225.5; DB 4; Length 276;

Best Local Similarity 28.5%; Pred. No. 4.3e-13;

Matches 63; Conservative 36; Mismatches 99; Indels 23; Gaps 5;

QY 243 GVSSTGG-----YETTKFVSMDLKPGQKVLVGGGCGGDFYMAETFDVYVQ 292

DB 35 GYVFDGEGQATFDENVMQMTDEMIRLDPADRDVLDIGCGGTGPAQLARDEYVVG 94

QY 293 FDLNVNMFALERS--IGLKCAVEFEVADCTKINYPNSDFVIYSRDTILHIQDKPALF 350

DB 95 TSVSARQVERGNRRBARAGLADRVRFQVDAMNLPDDGSGFDHCWALESLMHPDQOVL 154

QY 351 RSFYKMLPGGKVLISDYCKAGPPS--PFAAYIKQGY--DLHDVKEYQGMKDGAFYDV 408

DB 155 TEHVRVKEGARMPEIADVMVLYNPPDPSRPTATVSDTTIYAALTDIGDYDFIIRAAGMTVL 214

QY 409 LAEDTEQPIRVLRKELETVEKEKDVFTSDSEEDYNDIVG 449

DB 215 ELTDTITETAKTYGYVFWIRAHND-----EYVDIIG 246

RESULT 10

ABP96395

ID ABP96395 standard; protein; 280 AA.

XX ABP96395;

XX 22-MAY-2003 (first entry)

XX Nostoc punctiforme gamma-tocopherol methyltransferase SEQ ID NO:39.

XX Gamma-tocopherol methyltransferase; methyltransferase; tocopherol;
XX alpha-tocopherol; alpha-tocotrienol; stress; oxidative stress tolerance;
XX oxygen; ozone; UV tolerance; ultra violet tolerance; cold tolerance;
XX fungal pathogen; microbial pathogen; plant; enzyme.

XX Nostoc punctiforme.

XX WO2003016482-A2.

Db 5 LYQI-----QQFYDASSG-----LWEQIWEHMHGYGADGTOKKRRQAQIDL 50
 QY 253 TKEFYSLDLKPGQKVLNCGGIGGDFYMAETEDVEVVGFDLSVNMISFALERSI--GL 310
 Db 51 IEELNAGVQNAEDILVCGGIGGSSLYLAQFNATGTYLSPVQARATERALEANL 110
 QY 311 KCAVEFVADCTKINYPNSFDVYISDITLHQDKPALFRFPYKWLKPGKVLNIDYCK 370
 Db 111 SLRTOFQVANAQAFADSDLDVMSLESGEHMPDKTFLQBCYRVLKPGKGLNVTWCH 170
 QY 371 KAGPPSP 377
 Db 171 RPTDESP 177

RESULT 12
 AAB32507
 ID AAB32507 standard; protein; 283 AA.
 XX
 AC AAB32507;
 XX
 XX 19-JAN-2001 (first entry)
 XX
 XX S. lavendulae Mit M encoded protein sequence.
 XX
 XX Mitomycin; biosynthesis; mitosome ring system; antibiotic; anti-cancer;
 KW anti-inflammatory; immune-enhancer; immunosuppressant; asthma;
 KW chronic obstructive pulmonary disease; respiratory inflammation;
 KW fungicide; pesticide.
 XX
 XX Streptomyces lavendulae.
 XX
 XX WO200053737-A2.
 XX
 XX 14-SEP-2000.
 XX
 XX 10-MAR-2000; 2000WO-US006394.
 XX
 XX 12-MAR-1999; 98US-00266965.
 XX
 XX (MINU) UNIV MINNESOTA.
 XX (SHER/) SHERMAN D H.
 XX (MAOY/) MAO Y.
 XX (VARO/) VAROGLU M.
 XX (HEMM/) HE M.
 XX (SHEL/) SHELTON P C.
 XX
 XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon PC;
 XX
 XX WPI; 2000-601980/57.
 XX N-PSDB; AAC55802.
 XX
 XX Novel nucleic acid molecule comprising mitomycin biosynthetic gene
 XX cluster useful for cloning mitomycin biosynthetic genes for elucidating
 XX the molecular basis of mitosome ring system biosynthesis.
 XX
 XX Disclosure; Page 353; 399pp; English.
 XX
 XX This invention relates to isolated and purified nucleic acid molecules
 XX from the mitomycin biosynthetic gene cluster. Mitomycins are a group of
 XX natural products that contain a variety of functional groups, including
 XX amino benzocyclohexene and axiridine ring systems. The S. lavendulae
 XX mitomycin biosynthetic gene cluster comprises 47 mitomycin genes spanning
 XX 55kb of DNA. The invention includes an expression cassette comprising a
 XX mitomycin biosynthetic gene operably linked to a promoter, and host cells
 XX transformed with the cassette. The nucleotide, and protein sequences and
 XX the transformed host cells of the invention result in antiasthmatic,
 XX antiinflammatory, cytostatic, immunomodulatory, and antibiotic
 XX activities. The nucleotide sequences are used to elucidate the molecular
 XX basis for the biosynthesis of the mitosome ring system, as well as to
 XX engineer the biosynthesis of novel natural products, e.g. antibiotics,
 XX anti-inflammatory agents, anti-cancer agents, immune-enhancers,
 XX immunosuppressants, agents to treat asthma, chronic obstructive pulmonary

CC disease as well as other disease involving respiratory inflammation, or
 CC cholesterol-lowering agents or as crop protection agents (e.g. fungicides
 CC or insecticides) as well as biopolymers, e.g., in packaging of biomedical
 CC applications, or to engineer PHA monomer synthases. Sequences AAC55782-
 CC C55881, AAC55815-C55849 and AAB32485-B32542 represent mitomycin
 CC biosynthetic gene cluster DNA sequences and encoded proteins. Sequences
 CC AAC55812-C55814, AAC55850-C55856 and AAC55862-C55869 represent PCR
 CC primers used in the cloning of the mitomycin biosynthetic genes
 XX
 XX Sequence 283 AA;
 XX

Query Match 8.7%; Score 218.5; DB 3; Length 283;
 Best Local Similarity 26.9%; Pred. No. 2.2e-12;
 Matches 68; Conservative 46; Mismatches 92; Indels 47; Gaps 11;

QY 236 YERVFEGYVSTG-----GY-----ETTKFVSM-----LDLKPQKVLNCGGI 275
 Db 23 YDRFTALGNASLGENLHFGYWDSPDSQVPLAEYDDELTDMAERLRIGASVLDLGGV 82
 QY 276 GGGDFYMAETEDVEVVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPNSFDV 333
 Db 83 CTGPGVRIARLSGAHVGTGISVSHQVVRANALAEAGLADRARFQRADAMDLFFEDESFA 142
 QY 334 IYSRDTILHQDKPALFRFPYKWLKPGKVLNIDYCKAGPPSPFAAYIKQGYDLHD- 392
 Db 143 VIALESIIHMFDRQAQLAQRVLRPGRLVLTDFPERA-PLAPEGRAAVQRY---LHDF 198
 QY 393 -----VKEYGQMLKDG-----FYDVLADRTEQFIVLRKELETVKEKDVFIQDFSE 441
 Db 199 MMTVMVSAEAYPLLRGAGLWLEFLDI-SDQTEKTRLLS-----ERINSSKORLETQFG 254
 QY 442 E-----DYNDIVG 449
 Db 255 EMVNQFDPGLVG 267

RESULT 13
 ADE10279
 ID ADE10279 standard; protein; 283 AA.
 XX
 AC ADE10279;
 XX
 XX 29-JAN-2004 (first entry)
 XX
 XX S. lavendulae mitomycin biosynthetic protein MitM.
 XX
 XX Mitomycin biosynthetic protein; mitomycin C; antibiotic; MC; apoptosis;
 KW tumour hypoxia; cytostatic; anti-tumour agent; cancer.
 XX
 XX Streptomyces lavendulae.
 XX
 XX US2003134398-A1.
 XX
 XX 17-JUL-2003.
 XX
 XX 12-SEP-2001; 2001US-00953348.
 XX
 XX 12-SEP-2001; 2001US-00953348.
 XX
 XX (SHER/) SHERMAN D H.
 XX (MAOY/) MAO Y.
 XX (VARO/) VAROGLU M.
 XX (HEMM/) HE M.
 XX (SHEL/) SHELTON P.
 XX
 XX Sherman DH, Mao Y, Varoglu M, He M, Sheldon P;
 XX
 XX WPI; 2003-863498/80.
 XX N-PSDB; ADE10221.
 XX
 XX New nucleic acid molecule comprising a sequence having mitomycin
 XX biosynthetic gene cluster, useful for enhancing production of
 XX antibiotics.

KW cardiovascular disease; cancer; immune function; vitamin E;
KW alpha-tocopherol; photosynthetic organism;
KW tocopherol methyl transferase 2;
KW 2-methylphytylplastoquinol methyltransferase; tMT2; transgenic plant;
KW transgenic seed; mammalian diet; animal feed; feed source;
KW tocopherol biosynthesis; modified plant; oxidative stress tolerance;
KW UV tolerance; cold tolerance; fungal pathogen tolerance;
KW microbial pathogen tolerance; tMT2; enzyme; plant;
KW gamma-tocopherol transferase.
XX
XX
XX Anabaena sp.
XX OS
XX PN WO2003034812-A2.
XX XX
XX PD 01-MAY-2003.
XX XX
XX PF 24-OCT-2002; 2002WO-US034079.
XX XX
XX PR 25-OCT-2001; 2001US-0330563P.
XX XX
XX PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX PI Norris SR, Lincoln K, Stein JC, Valentin HE, Van Renennaam A;
XX XX
XX DR MPI; 2003-449223/42.
XX DR N-PSDB; ADD19104.
XX XX
XX PT New genes associated with the tocopherol biosynthesis, and its encoded
XX PT methyltransferase, useful for producing transgenic plants and seeds
XX PT expressing high tocopherol levels, which are useful as animal feed or
XX PT nutritional enhancer.
XX
XX PS Disclosure; SEQ ID NO 107; 243pp; English.
XX
XX
XX This invention relates to genes associated with the tocopherol
XX biosynthesis pathway. In particular, the invention relates to genes which
XX encode proteins with a methyltransferase activity. Tocopherols are an
XX important component in mammalian diets increased intake of which results
XX in a decreased risk of cardiovascular disease and cancer and improved
XX immune function. Vitamin E, or alpha-tocopherol, is significant for human
XX health and is primarily synthesised by plants and other photosynthetic
XX organisms. The present invention provides a nucleotide sequence which
XX encodes a tocopherol methyl transferase 2 (2-methylphytylplastoquinol
XX methyltransferase; tMT2) enzyme, which is a component of the tocopherol
XX biosynthesis pathway. The present invention is useful in methods of
XX reducing, in a plant, the expression of a gene encoding a plant
XX polypeptide having tMT2 activity, producing a plant having a seed with an
XX increased alpha-tocopherol level and producing a plant having a seed with
XX increased total tocopherol levels. The nucleic acid molecule is useful in
XX gene isolation, gene analysis, or in producing transgenic plants and
XX seeds that express high tocopherol or vitamin E levels. The plants are
XX useful in mammalian diet, particularly as animal feed, or for
XX nutritionally enhancing food and feed sources. The nucleic acid is
XX particularly useful for tocopherol biosynthesis in plants. The nucleic
XX acid molecule or polypeptide is also useful for producing modified plants
XX with improved tolerance to a variety of stress, for example oxidative
XX stress tolerance such as to oxygen or ozone, UV tolerance, cold
XX tolerance, or fungal/microbial pathogen tolerance. The present sequence
XX is the amino acid sequence of the Anabaena erecta gamma-tocopherol
XX methyltransferase enzyme related to the invention.

SQ Sequence 280 AA;

Query March 8.5%; Score 214; DB 7; Length 280;

Best Local Similarity 28.9%; Pred. No. 5.9e-12;

Matches 54; Conservative 35; Mismatches 68; Indels 30; Gaps 5;

QY 207 LMKQVSKDKGKQFRLDTSQYKCNLSILRYERFVGP-----GYVSTGYETTFEVSMLDL 262

DB.....5.LYQOI-----QQFYDASSG-----LWEIWEHMHGYYGADCTQKNERQAIJD 50

QY 263 -----KPGQKVLVGGGDFVMAETFDVEVGFDLNVNMFALERS--IGL 310

Db 51 ISELLTWAGVQTARNILDVCGGIGGSSLYLACKLNKAKATGTTLSFVQAAATERAKEAGL 110
QY 311 KCAVEFEVADCTKINYPDNSFOVIYSRDTILHTQDKPALFRSFYKWLKPGGKVLISDYCK 370
Db 111 SGRSQFLVANAQAAMPFDNSFDLVMSLESGEHPDKTKFLQBCYRVLKPGGKLIINVTTWCH 170
QY 371 KAGPPSP 377
Db 171 RPTDKTP 177

Search completed: July 26, 2004, 13:14:10
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:13:06 ; Search time 19 Seconds
(without alignments)
1285.215 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2515

Sequence: 1 HTVDLTETAMWLSQASDLID.....KURRTAKGEQRKGLFVAKKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgm2_6/ptodata/2/iaa/5A COMB.pdp:*
2: /cgm2_6/ptodata/2/iaa/5A COMB.pdp:*
3: /cgm2_6/ptodata/2/iaa/6A COMB.pdp:*
4: /cgm2_6/ptodata/2/iaa/6S COMB.pdp:*
5: /cgm2_6/ptodata/2/iaa/PCTRUS COMB.pdp:*
6: /cgm2_6/ptodata/2/iaa/backfileal.pdp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218.5	8.7	283	4	US-09-266-965-109
2	217.5	8.6	363	3	US-09-041-718-5
3	210.5	8.4	317	4	US-09-118-637A-2
4	194.5	7.7	283	3	US-09-036-987A-13
5	194.5	7.7	283	3	US-09-370-700-13
6	194.5	7.7	283	4	US-09-603-207-13
7	185.5	7.4	280	3	US-09-029-603-2
8	184.5	7.3	249	4	US-08-134-000C-6713
9	184.5	7.3	318	4	US-09-382-906A-2
10	180.5	7.2	275	4	US-09-266-965-110
11	177	7.0	115	4	US-09-266-965-6
12	173.5	6.9	348	4	US-09-118-637A-4
13	173	6.9	359	4	US-09-328-352-8005
14	172	6.8	275	3	US-09-036-987A-7
15	172	6.8	275	3	US-09-370-700-7
16	172	6.8	275	4	US-09-603-207-7
17	169.5	6.7	114	4	US-09-266-965-2
18	167.5	6.7	361	4	US-08-439-554-30
19	165	6.6	256	4	US-09-328-352-7702
20	163.5	6.5	361	3	US-09-041-718-4
21	161.5	6.4	382	4	US-09-644-907B-11
22	161	6.4	115	4	US-09-266-965-1
23	158	6.3	327	4	US-09-352-991A-28744
24	156.5	6.2	258	4	US-09-134-001C-3244
25	155.5	6.2	385	4	US-09-644-907B-2
26	154.5	6.1	295	4	US-09-540-236-3434
27	154.5	6.1	317	4	US-09-489-039A-8044

28	154.5	6.1	376	3	US-09-041-718-2
29	153	6.1	383	3	US-09-041-718-3
30	151	6.0	390	4	US-09-489-039A-9705
31	150	6.0	282	1	US-08-457-245-9
32	149.5	5.9	116	4	US-09-266-965-7
33	149	5.9	253	4	US-09-489-039A-8084
34	145.5	5.8	407	4	US-09-543-681A-4434
35	144.5	5.7	404	4	US-09-252-991A-19166
36	143.5	5.7	255	4	US-09-543-681A-5713
37	134	5.3	439	4	US-08-311-731A-48
38	131.5	5.2	260	4	US-09-489-039A-12651
39	129	5.1	271	4	US-09-252-991A-21522
40	126.5	5.0	298	4	US-09-644-907B-8
41	124	4.9	209	4	US-09-252-991A-32577
42	124	4.9	456	4	US-09-252-991A-25142
43	123	4.9	393	4	US-09-634-238-274
44	122	4.9	253	4	US-09-180-109A-35
45	122	4.9	254	4	US-09-180-109A-31

ALIGNMENTS

RESULT 1
US-09-266-965-109
; Sequence 109, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600,456US1
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 109
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Streptomyces lavendulae
US-09-266-965-109

Query Match 8.7%; Score 218.5; DB 4; Length 283;
Best Local Similarity 26.9%; Pred. No. 3.7e-15;
Matches 68; Conservative 46; Mismatches 92; Indels 47; Gaps 11;
QY 236 YERVFGEYVSTG-----GY-----ETTKEFVSM-----IDLKPGQKVLVDVCGI 275
DB 23 YDRFTALGAASLGENLHFGYWDSPDSQVPLAETDRLTDMMAERLRCAGSRVLDLACGV 82
QY 276 GGGDFYMAETFEVVGFDLSVNM-----SEALERSIGLKNVEFFVADCTKINYPDPSFDV 333
DB 83 GTPGVRRLSARHGVTGIVSHEQVVRANALAEAGLADRARFQRADAMDLPFEDESFDA 142
QY 334 IYSRDTLHQDPALFRSFYKWLKPGGKVLISDYCKKAGPPSPFAFYIKQGYDLDH- 392
DB 143 VIALESITHMPDRAQLAQVGRVLRPGRLVLTDFERA-PLAPEGRAAVQRY---LHDF 198
QY 393 -----VKEYGOMLKQAG-----FVDVLAEDTRQPIRVLRKELETVEKEKVFISDFSE 441
DB 139 MWTVWSAEAPPLRGNGLMLEFLDI-SDQLETKYFLLS-----ERINSSKQRIETQFGE 254
QY 442 E-----DYNDIVG 449

```

DB      255 EMVNPDPGLVG 267

RESULT 2
US-09-041-718-5
; Sequence 5, Application US/09041718A
; Patent No. 6225075
; GENERAL INFORMATION:
; APPLICANT: Bard, Martin
; TITLE OF INVENTION: DNA encoding sterol methyltransferase
; FILE REFERENCE: 740.003051
; CURRENT APPLICATION NUMBER: US/09/041,718A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Triticum atrivum
US-09-041-718-5

Query Match      8.6%; Score 217.5; DB 3; Length 363;
Best Local Similarity 29.8%; Pred. No. 7.2e-15;
Matches 71; Conservative 34; Mismatches 90; Indels 43; Gaps 8;

QY      254 KPFVSM-IDLKPGQKVLVGGCGGDPYMAETFDVVEVDFLNVNML--FALERSIGL 310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      108 EHFLALQLKELKPGMKVLVGGCGGCPLEIREIARFSSTVTGLANNVDYQITRGKALNRSVGL 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      311 KCANFEVADCTKINYPNSFDVYSDTTLIHQDKPALFRSFYKWLKPGKVLISDYC- 369
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      168 GATCDFVADFMKPPSNTDVAVATEATCHAPDPVGCYKRIYVRLKPGQCFAYVNCI 227
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      370 KNAGPSPEFAAYK--ORCYDLHDVKEYGQML--KDAGFVDV---LADR----- 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      228 THYDPNNATHKRIKDEILGNLPLDIRSTQCLOAVKADAGFEVVDKLAEDSPLPWYL 287
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      414 -----TEQFIRVLKLEETVEKVDVIFSDSEEDYNDIVGG 450
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      288 PLDPSPFSLSPFLTAVTGRIITRNKVL--EYVGLAPEGSRQVSSFLKAAEAGLVEG 343

RESULT 3
US-09-118-637A-2
; Sequence 2, Application US/09118637A
; Patent No. 6642434
; GENERAL INFORMATION:
; APPLICANT: DellaFenna, Dean
; APPLICANT: Shintani, David K.
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/118,637A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 920905.90024
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 317 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-118-637A-2

Query Match      8.4%; Score 210.5; DB 4; Length 317;
Best Local Similarity 26.4%; Pred. No. 3.4e-14;
Matches 55; Conservative 48; Mismatches 72; Indels 33; Gaps 7;

QY      236 YERVFGP-----GYVSTGGYETTKEFVSMIDL-----KPGQKVLVGGCGIGG 277
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      49 WEDVGEHMHGYYGPHGTVRIDRRQAQIDLIKELLAWAVPQNSAKP-RKILDLGCGGIGG 107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      278 GDFYMAETFDVVEVDFLNVNMLSPALE--RSIGLKCAVEFVADCTKINYPNSFDVIY 335
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      108 SLLYLAQHQAQVNGASLSPQVVERAGERARALGLSTCQFQVANALDLPASDSFDVWV 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      336 SRDTILIHQDKPALFRSFYKWLKPGKVLISDYCKA-----GPSPPEFAAYIKQGYDL 390
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      168 SLEGEHMPNKAQPLQEAWEVLKPGGRLILATWCHRPIDPGNGPLTADERRHL-QAIVDV 226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      391 H-----DVKEYGQMLKADAGFVDVLAED 412
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      227 YCLPVVSLPDYEAIARECGGGEIKTAD 254
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-09-036-987A-13
; Sequence 13, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Fatti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 283 amino acids

```

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-13

Query Match
Best Local Similarity 7.7%; Score 194.5; DB 3; Length 283;
Best Local Similarity 24.1%; Pred. No. 1.6e-12;
Matches 57; Conservative 47; Mismatches 106; Indels 27; Gaps 6;
234 LAYERVFGVGVSTGGY-----ETTKFVSMCLKPGQKVLDPGCGIGGGD 279
8 LAHGRPLHGYWA--GGYREDAGATPMSDAADQLTDLFIDKALRPGAHFLDLGCGNGQPV 66
280 FYMAETFEVVGFDLSVNMISFA--LERSIGLKCAVEFEVADCTKINYPNSPDVIYSR 337
67 VRAACASGVRVTGTVNAQHLAAATRLANETGLAGSLFPLVDGAGLPYDFGFFQAAWAM 126
338 DTILHIQDKPALFRFYKWLKPGQKVLISDYCKKAGPPSPFAAYIKORGYDLHDVKEYG 397
127 QSVVQIVDQAAAIREVHRLPGGRFVLGDIITRVRLPE--EYAAV--WTGTTAHTLNSFT 183
398 QMLKDGAFVDVLAEDRTEQ-----FIRVLRKELETVEKEKDVPSFSEEDYNDI 447
184 ALVSEAGFEILEVTDLTATQTCMVSWYVDLLAKDLDELACGVEPAAVGTQQRYLGDI 240

RESULT 5
US-09-370-700-13
; Sequence 13, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIVI
; CURRENT APPLICATION NUMBER: US/09/370,700
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; TYPE: PRT
; LENGTH: 283
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-13

Query Match
Best Local Similarity 7.7%; Score 194.5; DB 3; Length 283;
Best Local Similarity 24.1%; Pred. No. 1.6e-12;
Matches 57; Conservative 47; Mismatches 106; Indels 27; Gaps 6;
234 LAYERVFGVGVSTGGY-----ETTKFVSMCLKPGQKVLDPGCGIGGGD 279
8 LAHGRPLHGYWA--GGYREDAGATPMSDAADQLTDLFIDKALRPGAHFLDLGCGNGQPV 66
280 FYMAETFEVVGFDLSVNMISFA--LERSIGLKCAVEFEVADCTKINYPNSPDVIYSR 337
67 VRAACASGVRVTGTVNAQHLAAATRLANETGLAGSLFPLVDGAGLPYDFGFFQAAWAM 126
338 DTILHIQDKPALFRFYKWLKPGQKVLISDYCKKAGPPSPFAAYIKORGYDLHDVKEYG 397
127 QSVVQIVDQAAAIREVHRLPGGRFVLGDIITRVRLPE--EYAAV--WTGTTAHTLNSFT 183
398 QMLKDGAFVDVLAEDRTEQ-----FIRVLRKELETVEKEKDVPSFSEEDYNDI 447
184 ALVSEAGFEILEVTDLTATQTCMVSWYVDLLAKDLDELACGVEPAAVGTQQRYLGDI 240

RESULT 6
US-09-603-207-13
; Sequence 13, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIVI
; CURRENT APPLICATION NUMBER: US/09/603,207B
; EARLIER FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-13

Query Match
Best Local Similarity 7.7%; Score 194.5; DB 4; Length 283;
Best Local Similarity 24.1%; Pred. No. 1.6e-12;
Matches 57; Conservative 47; Mismatches 106; Indels 27; Gaps 6;
234 LAYERVFGVGVSTGGY-----ETTKFVSMCLKPGQKVLDPGCGIGGGD 279
8 LAHGRPLHGYWA--GGYREDAGATPMSDAADQLTDLFIDKALRPGAHFLDLGCGNGQPV 66
280 FYMAETFEVVGFDLSVNMISFA--LERSIGLKCAVEFEVADCTKINYPNSPDVIYSR 337
67 VRAACASGVRVTGTVNAQHLAAATRLANETGLAGSLFPLVDGAGLPYDFGFFQAAWAM 126
338 DTILHIQDKPALFRFYKWLKPGQKVLISDYCKKAGPPSPFAAYIKORGYDLHDVKEYG 397
127 QSVVQIVDQAAAIREVHRLPGGRFVLGDIITRVRLPE--EYAAV--WTGTTAHTLNSFT 183
398 QMLKDGAFVDVLAEDRTEQ-----FIRVLRKELETVEKEKDVPSFSEEDYNDI 447
184 ALVSEAGFEILEVTDLTATQTCMVSWYVDLLAKDLDELACGVEPAAVGTQQRYLGDI 240

RESULT 7
US-09-029-603-2
; Sequence 2, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospiech, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029,603
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Streptomyces longisporoflavus
; FEATURES:
; OTHER INFORMATION: methyl transferase-like protein
US-09-029-603-2

```

Query Match      7.4%; Score 185.5; DB 3; Length 280;
Best Local Similarity 26.1%; Pred. No. 1.5e-11;
Matches 68; Conservative 42; Mismatches 92; Indels 59; Gaps 13;

QY 236 YERY-----FGP-----GYSTGGYETIKK-----FVSMDLKPGQKVLNVC 273
DB 18 YDLTSLAMNDGSPNVHIGYWDTPGSEATIEEAMDRITDVFIERLNAYATSHVLDGC 77
QY 274 GIGGGFMAETDVEVGGFSLVNMISPA--LERSIGLKCAVEPEVADCTKINYPNSF 331
DB 78 GVGGFGLRVARTGARTGISIEEQIRANLAAEAGVADRAVQHGDKMLPFADASF 137
QY 332 DVIVSRDTHIQKPALFRSFKMLKPGKVLISDYCKKAGPPSPFAANYIKORYD-- 389
DB 138 DAVMALESICHMPDROQVTEVCVRVLRPGGRIVLTDIFER-----HPRKA--VRHFGIDKF 191
QY 390 -----LHDVKEYGQMLKDG-----FVDVLAEDRTEQFIRVLKR--LETVEKRDV 434
DB 192 CRDLMTTADIDVALLHRSGLRLRIVDV-----TEQTLRLADEIGRLAAVE-ERP 245
QY 435 FISO-----FSEEDV--NDIVG 449
DB 246 ANDGNGFAGDDSPKPSLAG 266

```

```

RESULT 8
US-09-134-000C-6713
; Sequence 6713, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134.000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6713
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6713

```

```

Query Match      7.3%; Score 184.5; DB 4; Length 249;
Best Local Similarity 25.6%; Pred. No. 1.6e-11;
Matches 72; Conservative 49; Mismatches 109; Indels 51; Gaps 12;

QY 209 QKVDSKDDK-----GQRFPLDTQYKNSILRYRVRPGVSTGGYETTKKPFVSMDLK 263
DB 2 RRVQMKENYKDDNIFQKYQSRSQ-----KGLAGAWETLKKWLP--DPK 47
QY 264 PQKQVLDVGGGIGGGDFMAETDVEVGGFSLVNMISPALERSIGLKCAVEPEVADCTK 323
DB 48 -GRVLDLGGYGMWCIYAMENGASSVVGVDISHKMLKAVAKTHFPQ--IETECCAIED 104
QY 324 INTPDMSFVIVSRDTHIQKPALFRSFKMLKPGKVLISDYCKKAGPPSPFAA-- 381
DB 105 VDPPESEFVILLSSLAFFVADYENLKKIYRMLKAGGNLVF-----VEHFPVTANG 157
QY 382 ----YIKQGYDLHDVKE--YQMLKDGAFVDVLAEDRTEQFIRVLKRLETVEKRDV- 434
DB 158 TDQWYNEKEILLHPVDVNYEGKRTAMFL-----EKKVYKHTLTLYLNTLLSNFII 213
QY 435 --FISDFSEEDYNDIVGGWMDKLRRTAKGQRMGLFWAKK 473
DB 214 NQIVEPQPPENMDI-PGMADEKRRPMM-----LIVSACK 247

```

```

RESULT 9
US-09-382-906A-2

```

```

; Sequence 2, Application US/09382906A
; Patent No. 6448475
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Spintani, David
; TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants
; FILE REFERENCE: 920905.90032
; CURRENT APPLICATION NUMBER: US/09/382,906A
; CURRENT FILING DATE: 1999-08-25
; PRIOR FILING DATE: 1999-08-25
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Synecchocystis PCC6803
US-09-382-906A-2

```

```

Query Match      7.3%; Score 184.5; DB 4; Length 318;
Best Local Similarity 29.6%; Pred. No. 2.3e-11;
Matches 47; Conservative 30; Mismatches 61; Indels 21; Gaps 5;

QY 227 QYKNSILRYRVRPGVSTGGY--ETTKGFV-SMLD-----LKPQKVL 269
DB 42 QWTEGDILEY---YMGDHIHLGHYGGPPVAKDFIQSKIDFVHAMAQWGLDLPPTGTVL 98
QY 270 DVGGGIGGGDFMAETDVEVGGFSLVNMISPALERSIGLKCAVEPEVADCTKINYPDN 329
DB 99 DVGGGIGGGSRILAKDYGFNVTTITISPOQVKEATLT--PPDVTAKFAVDDAMALSPFDG 157
QY 330 SPDVIVSRDTHIQKPALFRSFKMLKPGKVLISDY 368
DB 158 SPDVVSVSEAGPHFDPKAVFAKELLRVVKGGLVVDW 196

```

```

RESULT 10
US-09-266-965-110
; Sequence 110, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.45GUSI
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 110
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Streptomyces lavendulae
US-09-266-965-110

```

```

Query Match      7.2%; Score 180.5; DB 4; Length 275;
Best Local Similarity 26.2%; Pred. No. 5e-11;
Matches 59; Conservative 39; Mismatches 105; Indels 19; Gaps 7;

QY 242 PGVSTGGYETTKKPFVSMDLKPGKVLISDYCKKAGPPSPFAAETDVEVGGFSLVNMIS 301
DB 41 PTVVEAGRLTDYVSRRLRPAQGERVLDVSSGKATLRAARHVRATG--VSNPYQ 98
QY 302 PALERSIGLK--CAVEPEVADCTKINYPDMSFDVIVSRDTHIQKPALFRSFKMLK 358

```


Db 99 VCLSRQAEGBDEATEFRIGOMLAPFPDGSFDACYAESICHALRADVFTETARVLR 158
QY 359 PGKVLISDYCKKAGPSPFAFYIKORGVDLHDVKEYGOMLKDAGFVDVLAEDRTE--Q 416
Db 159 FGRVTVTDVFLRR--PLSDASRTIVDTA---NDNFQQQPVLTREAYEDCHRSVGLWEVE 213
QY 417 FIRV---LRELETVKE---KDVFSDFSEEDYNDIVGG 450
Db 214 FLDIGDEVRSYEAAKWAARDELGSHMDDEAFERWVDG 254

RESULT 11

US-09-266-965-6
; Sequence 6, Application US/09266965
; Patent No. 6495348
; GENERAL INFORMATION:
; APPLICANT: Sherman, D
; APPLICANT: Mao, Y
; APPLICANT: Varoglu, M
; APPLICANT: He, M
; APPLICANT: Sheldon, P
; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
; FILE REFERENCE: 600.456U51
; CURRENT APPLICATION NUMBER: US/09/266,965
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 08/624,447
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: PCT/US94/11279
; EARLIER FILING DATE: 1994-10-06
; EARLIER APPLICATION NUMBER: US 08/133,963
; EARLIER FILING DATE: 1993-10-07
; SOFTWARE: PASTSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Amycolatopsis mediterranei
US-09-266-965-6

Query Match 7.0%; Score 177; DB 4; Length 115;

Best Local Similarity 33.9%; Pred. No. 2.9e-11;
Matches 37; Conservative 26; Mismatches 44; Indels 2; Gaps 1;

QY 262 LKPGQKVDVCGGIGGGDFYMAETFDVEVVGFDLSVNMISFALERS--IGLKCAVEFEVA 319
Db 2 LRAGDRLLDGGCGEPAIRMATANDVMVTGISSEKQVERANDRAYKADVDVDFVFEYA 61
QY 320 DCTKINYDPSFDVYSDTILHIQDKPALFRSFYKWLKPGGKVLISDY 368
Db 62 DANELFPYDASPDVVWVALESLSHMPDRHWVIRQARVLRPGGRLALGDF 110

RESULT 12

US-09-118-637A-4
; Sequence 4, Application US/09118637A
; Patent No. 6642434
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; APPLICANT: Shintani, David K.
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patencin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/118,637A

FILING DATE:
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 920905.90024

TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-118-637A-4

Query Match 6.9%; Score 173.5; DB 4; Length 348;
Best Local Similarity 27.2%; Pred. No. 4.3e-10;
Matches 52; Conservative 30; Mismatches 64; Indels 45; Gaps 7;

QY 202 NQISLWQKV-----DSK---DDKGFQ-----RFLDTQYKCNLSILRYRVEGP 242
Db 69 NETSGLMEEIWDHMHGFDYDPSVQLSDSGHKEAQIRMIIES-----LRPAGV--- 118

QY 243 GYVSTGGYVETKESFVSMILDKPGQKVDVCGGIGGGDFYMAETFDVEVVGFDLSVNMISF 302
Db 119 -----TDEE-----EKKIKVVDVCGGIGSSRYLASKFGACIGITLSPVQAKR 164

QY 303 A--LERSIGLKCAVEFEVADCTKINYDPSFDVYSDTILHIQDKPALFRSFYKWLKPG 360
Db 165 ANDLAAQSLSHKASFOVADALDQPFEDGKFDLWSMESGEHMPDKAKFVKELVRVAAPG 224

QY 361 GKVLISDYCKK 371
Db 225 GRILITVWCHR 235

RESULT 13
US-09-328-352-8005

; Sequence 8005, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

; APPLICANT: Gary L. Breston et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8005

; LENGTH: 359
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8005

Query Match 6.9%; Score 173; DB 4; Length 359;
Best Local Similarity 25.6%; Pred. No. 5.2e-10;
Matches 69; Conservative 48; Mismatches 95; Indels 58; Gaps 12;

QY 195 VRNKKQHQIISWLMQKVDKDDKFGFRLDTQYKCNLSILRYRVEFGYVGTGYYETK 254
Db 124 VRTEKAKVAEVEFHSVASKYD-----LNDLDM-----SFGHRLWK 160

QY 255 EF-VSMLDLKPQKVDVCGGIGGGDFYMAETFDVE-----VVGFDLSVNMISFALERS 307
Db 161 REAINNSGVRRGQHVLDIAGTGD-----LAKVFSREVGQGHVLSLINESMLNVRDRL 216

QY 308 IGLKCA-VEFEVADCTKIN-YPDNSFDVYSDTILHIQDKPALFRSFYKWLKPGGKVL 365

Db 217 IDAGTNNVDFVLANAETLEPADNSFDLVTFISGLRNVTDKDAALASMPFVLKPGGRLV 276
Qy 366 SDYCKAGPPSPFAAIVKQGYDLHDVKEYGQMLK---DAGFVDVLAED-RTEQFIRV 420
Db 277 LEFSK-----PVPEPSKL--YDLSPTALPIMGKLVANDSESVKYLAESIRNHPDPT 328
Qy 421 LRKELETVEKEKQVPSDFSEEDYNDIVGG 450
Db 329 LKGMWEN-----AGFONCDYHNLITGG 349

RESULT 14
US-09-036-987A-7
; Sequence 7, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 275 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-7

Query Match 6.8%; Score 172; DB 3; Length 275;
Best Local Similarity 23.3%; Pred. No. 4.3e-10;
Matches 55; Conservative 39; Mismatches 102; Indels 40; Gaps 6;
Qy 241 GP-----GYVSTGGYETTKFVSMLD-----LKPQKVLVDVGGIGGGDFTYMAETFD 287
Db 29 GPCTAHGGYWENDGRASWQQAADRLTDLVAERTVLDGGVRLDVGCGTGQPALRVARNA 88
Qy 288 VEVVGFGLSVNMISFALB--RSIGLKCAVEFVADCTKINYPDNSFDVIYSDTILHIQD 345
Db 89 IQTGTISQVQVAIAADCAERGLSHRVDFSCVDAMSIFYPDNAPDAAMQSLLENSE 148
Qy 346 KPALFRSFYKWLKPGKVLISDYCKKAGPPSPFAAIVKQGYDLHDVKEYGQMLKADGF 405

Db 149 PDRAIREILRVLPKGGILGVTVYVVKREAGGMPVSGDRMPTGLRICLAELQLLESLRAAGF 208
Qy 406 VDVLAEEDRTQFIRVLKKELETVEKEKQVPSDFSEE-----DYNDIVGGW 451
Db 209 -EIL-----DWEDEVSSRTRVFMPPQFAEELAAHQHGIADRYGPAVAGM 249
RESULT 15
US-09-370-700-7
; Sequence 7, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 275
; TYPE: PNT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-7

Query Match 6.8%; Score 172; DB 3; Length 275;
Best Local Similarity 23.3%; Pred. No. 4.3e-10;
Matches 55; Conservative 39; Mismatches 102; Indels 40; Gaps 6;
Qy 241 GP-----GYVSTGGYETTKFVSMLD-----LKPQKVLVDVGGIGGGDFTYMAETFD 287
Db 29 GPCTAHGGYWENDGRASWQQAADRLTDLVAERTVLDGGVRLDVGCGTGQPALRVARNA 88
Qy 288 VEVVGFGLSVNMISFALB--RSIGLKCAVEFVADCTKINYPDNSFDVIYSDTILHIQD 345
Db 89 IQTGTISQVQVAIAADCAERGLSHRVDFSCVDAMSIFYPDNAPDAAMQSLLENSE 148
Qy 346 KPALFRSFYKWLKPGKVLISDYCKKAGPPSPFAAIVKQGYDLHDVKEYGQMLKADGF 405
Db 149 PDRAIREILRVLPKGGILGVTVYVVKREAGGMPVSGDRMPTGLRICLAELQLLESLRAAGF 208
Qy 406 VDVLAEEDRTQFIRVLKKELETVEKEKQVPSDFSEE-----DYNDIVGGW 451
Db 209 -EIL-----DWEDEVSSRTRVFMPPQFAEELAAHQHGIADRYGPAVAGM 249

Search completed: July 26, 2004, 13:16:23
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:15:32 ; Search time 46 Seconds
(without alignments)
3220.042 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2515

Sequence: 1 HVVDLTTEAMMLDSQASDLDD.....KLRTAKGEQRWGLFVAKKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2011	80.0	513	12	US-10-425-114-37470 Sequence 37470, A
2	1932.5	76.8	530	16	US-10-437-963-155711 Sequence 155711, A
3	1829.5	72.7	491	14	US-10-149-759-82 Sequence 82, Appl
4	1760	70.0	501	12	US-10-424-599-172601 Sequence 172601, A
5	1654	65.8	399	12	US-10-425-114-72568 Sequence 72568, A
6	1321	52.5	314	12	US-10-425-114-69972 Sequence 69972, A
7	1088	43.3	331	12	US-10-425-114-44987 Sequence 44987, A
8	1064	42.3	377	12	US-10-425-114-60338 Sequence 60338, A
9	937	37.3	219	12	US-10-425-114-53753 Sequence 53753, A
10	839	33.4	237	12	US-10-424-599-257495 Sequence 257495, A
11	690.5	27.5	271	16	US-10-437-963-110035 Sequence 110035, A
12	665.5	26.5	183	12	US-10-424-599-240673 Sequence 240673, A
13	658	26.2	168	14	US-10-149-759-50 Sequence 50, Appl
14	516.5	20.5	437	15	US-10-369-493-6306 Sequence 6306, Ap
15	516.5	20.5	437	16	US-10-602-268-21 Sequence 21, Appl

15	496.5	19.7	472	16	US-10-602-268-12	Sequence 12, Appl
17	480	19.1	437	16	US-10-602-268-11	Sequence 11, Appl
18	425	16.9	161	12	US-10-424-599-224095	Sequence 224095, A
19	421	16.7	133	16	US-10-437-963-110034	Sequence 110034, A
20	413	16.4	475	16	US-10-602-268-19	Sequence 19, Appl
21	413	16.4	484	16	US-10-602-268-20	Sequence 20, Appl
22	404	16.1	460	16	US-10-602-268-7	Sequence 7, Appl
23	399	15.9	460	16	US-10-602-268-8	Sequence 8, Appl
24	396	15.7	457	16	US-10-602-268-9	Sequence 9, Appl
25	375	14.9	120	16	US-10-437-963-155679	Sequence 155679, A
26	353.5	14.1	469	16	US-10-602-268-10	Sequence 10, Appl
27	296	11.8	155	12	US-10-424-599-216620	Sequence 216620, A
28	286.5	11.5	264	15	US-10-369-493-12004	Sequence 12004, A
29	245	9.7	111	16	US-10-437-963-149311	Sequence 149311, A
30	243.5	9.7	117	12	US-10-424-599-174519	Sequence 174519, A
31	241.5	9.6	281	14	US-10-132-134-20	Sequence 20, Appl
32	224	8.9	280	14	US-10-279-029-106	Sequence 106, Appl
33	224	8.9	280	14	US-10-219-810-39	Sequence 39, Appl
34	218.5	8.7	283	10	US-09-953-348-109	Sequence 109, Appl
35	218.5	8.7	283	14	US-10-267-255-109	Sequence 109, Appl
36	214	8.5	280	14	US-10-279-029-107	Sequence 107, Appl
37	214	8.5	280	14	US-10-219-810-40	Sequence 40, Appl
38	214	8.5	286	15	US-10-369-493-18938	Sequence 18938, A
39	210.5	8.4	317	14	US-10-219-810-41	Sequence 41, Appl
40	201.5	8.0	243	15	US-10-369-493-3806	Sequence 3806, Ap
41	201.5	8.0	362	16	US-10-437-963-147768	Sequence 147768, A
42	200.5	8.0	345	14	US-10-279-029-95	Sequence 95, Appl
43	200.5	8.0	345	14	US-10-219-810-22	Sequence 22, Appl
44	199	7.9	344	9	US-09-779-144A-7	Sequence 7, Appl
45	198	7.9	338	12	US-10-425-114-41948	Sequence 41948, A

ALIGNMENTS

RESULT 1

US-10-425-114-37470
; Sequence 37470, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 37470
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; OTHER INFORMATION: Clone ID: LIB22-074-F3_FLI.pep
US-10-425-114-37470

Query Match	80.0%	Score	2011	DB	12	Length	513
Best Local Similarity	77.8%	Pred. No.	9.4e+189				
Matches	367	Conservative	51	Mismatches	54	Indels	0
						Gaps	0
QY	1	HVVDLTTEAMMLDSQASDLDDKEERPEILSMPLPEKCKLLELGAGIGRTGLAKAGQV	60				
Db	41	HSADLTVEAMMLDSRASDLDDKEERPEVLSLLPYEKSVLLELGAGIGRTGLAKAGEL	100				
QY	61	IALDPIEAIKKNVINGHYKNVKFKMCADVTSTPLTFPHSLDVFISNWLMLYLSDEEVE	120				
Db	101	IALDPIFNVKNESINGHYKNVKFKMCADVTSPDLATDGLDLSLIFSNWLMLYLSDEEVE	160				
QY	121	NLVERMLKWLPGGYIFPESCFHQSGDHKKSNPTHYREPPYTKAFKCHLQDQSGNS	180				

Db 161 LLAEWGMKVGSIYFFRSCPHQSDSKRKNPHTYRPFYSKVFQECOTRDAAGNS 220
 Qy 181 YELSLCKCIGAVRKNQKQIISWLMQVSKDDKGPORLDTSOYKCHSILRYERVF 240
 Db 221 FELSMIGCKCIGAVRKNQKQIISWLMQVSKDDKGPORLDTSOYKCHSILRYERVF 280
 Qy 241 GPGVSTGGYTTTKEFVSMMLDLKPGQKVLVDGCGIGGDPFYMAETFDVEVGVGDLVSNNI 300
 Db 281 GCGFVSTGGYTTTKEFVSMMLDLKPGQKVLVDGCGIGGDPFYMAETFDVEVGVGDLVSNNI 340
 Qy 301 SPALERSIGLCAVEFEVADCTKINYPDNSFDVIYSRDTILHIDQKPAFLFYSFKWLPKG 360
 Db 341 SPALERAIGLSCSVEFEVADCTKINYPDNSFDVIYSRDTILHIDQKPAFLFYSFKWLPKG 400
 Qy 361 GKVLISDYCKKAGPSPFPAYIKORGYDLHDVKEVGMKLDAGFVDVLAEDRTQPIRV 420
 Db 401 GKVLISDYCKKAGPSPFPAYIKORGYDLHDVKEVGMKLDAGFVDVLAEDRTQPIRV 460
 Qy 421 LRKELETVEKEDVIFSDPSEEDYNDIVGGMNDKLRRTAKGEQRWGLFVAKK 472
 Db 461 LRKELEVRKKEKFIISDFSKEDYDDIVGGMNDKLRRTAKGEQRWGLFVAKK 512

RESULT 2
 US-10-437-963-155711
 ; Sequence 155711, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Bouharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 155711
 ; LENGTH: 530
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MKT4530_5544C.1.pap
 US-10-437-963-155711

Query Match 76.8%; Score 1932.5; DB 16; Length 530;
 Best Local Similarity 71.4%; Pred. No. 5.3e-181;
 Matches 359; Conservative 52; Mismatches 61; Indels 31; Gaps 2;

Qy 1 HTVDTLEAMWLDQSASDLKKEPRLSMLPPLLEGKLELGLAGIGRFTGELAKAGOV 60
 Db 28 HSKDLTVAMWLDQSASDLKKEPRLSMLPPLLEGKLELGLAGIGRFTGELAKAGOV 60
 Qy 40 LELGAGIGRFTGELAKAGOVIALDFTSIAIKNVEVINGHYKVKPMCAVTSPTLSPP 99
 Db 88 LELGAGIGRFTGELAKAGOVIALDFTSIAIKNVEVINGHYKVKPMCAVTSPTLSPP 99
 Qy 100 HSLDVPISNMLMYSDEEVEKLVKGMVWLVKGVIFRSCPHQSDSKRKNPHTYR 159
 Db 148 NSIDLIFSNMLMYSDEEVEKLVKGMVWLVKGVIFRSCPHQSDSKRKNPHTYR 207
 Qy 160 EPRYTKAFTECHLDQSGNSYELSLCKCIGAVRKNQKQIISWLMQVSKDDKGP 219
 Db 208 EPRYTKAFTECHLDQSGNSYELSLCKCIGAVRKNQKQIISWLMQVSKDDKGP 267
 Qy 220 QRFLDTSQYKCHSILRYERVFPGYVSTGGY-----ETTKFVSMMLDLKPGQKVL 269

Db 268 QRFLDTSQYKCHSILRYERVFPGYVSTGGYVCLFFYLSPETTKFVDRLDLKPGQKVL 327
 Qy 270 DWGCGIGGDPFYMAETFDVEVGVGDLVSNNISFALERSIGLCAVEFEVADCTKINYPDN 329
 Db 328 DWGCGIGGDPFYMAETFDVEVGVGDLVSNNISFALERSIGLCAVEFEVADCTKINYPDN 387
 Qy 330 SPDVYSRDTILHIDQKPAFLFYSFKWLPKGQKVLISDYCKKAGPSPFPAYIKORGYD 389
 Db 388 TFDVYSRDTILHIDQKPAFLFYSFKWLPKGQKVLISDYCKKAGPSPFPAYIKORGYD 447
 Qy 390 LHDVKEVGMKLDAGFVDVLAEDRTQPIRVLAKELETVEKEDVIFSDPSEEDYNDIVG 449
 Db 448 LHDVKEVGMKLDAGFVDVLAEDRTQPIRVLAKELETVEKEDVIFSDPSEEDYNDIVG 507
 Qy 450 GNMNKLRTAKGEQRWGLFVAKK 472
 Db 508 GNMNKLRTAKGEQRWGLFVAKK 530

RESULT 3
 US-10-149-759-82
 ; Sequence 82, Application US/10149759
 ; Publication No. US20030157592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lerchl, Jens
 ; APPLICANT: Renz, Andreas
 ; APPLICANT: Ehrhardt, Thomas
 ; APPLICANT: Reinold, Andreas
 ; APPLICANT: Cirpus, Petra
 ; TITLE OF INVENTION: Mosa Genes from Physcomitrella patens encoding proteins
 ; FILE OF INVENTION: Involved in the synthesis of tocopherols and
 ; FILE REFERENCE: BASF/NAB 1333/99 PCT/US
 ; CURRENT APPLICATION NUMBER: US/10/149,759
 ; CURRENT FILING DATE: 2002-10-17
 ; PRIOR APPLICATION NUMBER: PCT/EP/00/12698
 ; PRIOR FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: WordPerfect version 6.1
 ; SEQ ID NO 82
 ; TYPE: PRT
 ; ORGANISM: Physcomitrella patens
 ; OTHER INFORMATION: Physcomitrella patens
 US-10-149-759-82

Query Match 72.7%; Score 1829.5; DB 14; Length 491;
 Best Local Similarity 70.0%; Pred. No. 6.6e-171;
 Matches 332; Conservative 65; Mismatches 74; Indels 3; Gaps 2;

Qy 1 HTVDTLEAMWLDQSASDLKKEPRLSMLPPLLEGKLELGLAGIGRFTGELAKAGOV 60
 Db 17 HSEVSEVAMWLDQSASDLKKEPRLSMLPPLLEGKLELGLAGIGRFTGELAKAGOV 76
 Qy 61 IALDFTIESAIKKNVEVINGHYKVKPMCAVTSPTLSPPHSLDLVIFSNMLMYSDEEVE 120
 Db 77 LAMPFMENTIKKEDVNGHYNNIDFKCADVTSPLDNTAAGSADLVFNSMLMYSDEEVE 136
 Qy 121 NLVERMLKMPGGYIFFRSCPHQSDSKRKNPHTYRPFYSKVFQECOTRDAAGNS 180
 Db 137 GLASRVWMLRFGYIFFRSCPHQSDSKRKNPHTYRPFYSKVFQECOTRDAAGNS 195
 Qy 181 YELSLCKCIGAVRKNQKQIISWLMQVSKDDKGPORLDTSOYKCHSILRYERVF 238
 Db 196 FRFEMGCKCIGAVRKNQKQIISWLMQVSKDDKGPORLDTSOYKCHSILRYERVF 255
 Qy 239 VFGPGYVSTGGYTTTKEFVSMMLDLKPGQKVLVDGCGIGGDPFYMAETFDVEVGVGDLVSNNI 298
 Db 256 IFBGGFVSTGGYTTTKEFVSMMLDLKPGQKVLVDGCGIGGDPFYMAETFDVEVGVGDLVSNNI 315
 Qy 299 MISFALERSIGLCAVEFEVADCTKINYPDNSFDVIYSRDTILHIDQKPAFLFYSFKWLPKG 358
 Db 316 MISFALERSIGLCAVEFEVADCTKINYPDNSFDVIYSRDTILHIDQKPAFLFYSFKWLPKG 375

QY 359 PGKVLISDYCKKAGPPSPFAAYIKORGYDLHDVKEYGOMLKADGAFVDVLAEDRTQFI 418
 DB 376 PGKVLISDYCRAPQPSFAEAYIQORGYDLHDVKEYGOMLKADGAFVDVLAEDRTQFI 435
 QY 419 RVLRKELETVEKEKDVIFSDPSEEDYNDIVGWNKDLARTAKGEORWGLFVAKK 472
 DB 436 EVLQRELATTEAGRDQFINDFSEEDYNYIVSGMKSLKRCNSDEQKWLFIAYK 489

RESULT 4
 US-10-424-599-172601
 ; Sequence 172601, Application US/10424599
 ; Publication No. US20040031072AL
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 172601
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(501)
 ; OTHER INFORMATION: unsure at all xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_126075C.1.pap
 US-10-424-599-172601

Query Match 70.0%; Score 1760; DB 12; Length 501;
 Best Local Similarity 68.1%; Pred. No. 4.7e-164;
 Matches 323; Conservative 68; Mismatches 81; Indels 2; Gaps 1;

QY 1 HTVDITIAWMLDSQADLKEPERITLSMLPLECKLLELGAIGRTPTGLAEKAGQV 60
 DB 27 HSADLSVEMWMLDSKAHLDEERPEVLSLLPPEKSGVLELGLNGIPTGELAKAQL 86
 QY 61 IALDIESAIKNEVINGHYKNKPMCAVTSPTLS--PPHSLDVIISNMLMLSLDER 118
 DB 87 LAVDFIDTAIKNETINGHHNVKFLCADVTSNNNNVSEGSVOVFNKXLMILYLSID 146
 QY 119 VENLVERMLKWLKPGGYIFFRSCFHQSGDHKRSNPTHYRPFYTKAFKCHLODGS 178
 DB 147 VEKFDTRMIRLLIDGGYVFFKGTCTQSGNSKRSKYNSTHYKVPFNTKVIKCHMSDNKG 206
 QY 179 NSVELSLSCKICIGAYVRNKKONQISWLMQKVDSDKDGKFORFLDTSQYKNSILRYER 238
 DB 207 NSFELSLVGCCKICIGAYVRNKKONQICWISLVRSDQDRGFQRLDRVYSHKSLRYER 266
 QY 239 VFGPGVSTGGYTTKGFVPMMLDKPGQKVLVDCGCGIGGDFYMAFTFVVGFDLSVN 298
 DB 267 MTGPGFVSTGGLETTKTFKFAVGLKPGQKVLVDCGCGIGGDFYMAENFVVGIGDLSIN 326
 QY 299 MISFALERSIGLKAVEPADCTKINYPNPSDVIYSRDTILHIQOKPALFRSPFVWLK 358
 DB 327 MISLAIERAIGLNAVEFDCACVRYTPENTFVITRDTMLHVKKDTLFRSPFVWLK 386
 QY 359 PGKVLISDYCKKAGPPSPFAAYIKORGYDLHDVKEYGOMLKADGAFVDVLAEDRTQFI 418
 DB 387 PGKVLISDYCKKAGPPSPFAAYIKORGYDLHDVKEYGOMLKADGAFVDVLAEDRTQFI 446
 QY 419 RVLRKELETVEKEKDVIFSDPSEEDYNDIVGWNKDLARTAKGEORWGLFVAKK 472
 DB 447 NTLQELNALENKODFDIGDFSEEDYNEIVERWKAQTRCASDEQKWLFIAYK 500

RESULT 5
 US-10-425-114-72568
 ; Sequence 72568, Application US/10425114
 ; Publication No. US20040034888AL
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 72568
 ; LENGTH: 399
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB23-061-D6_FLI.pap
 US-10-425-114-72568

Query Match 65.8%; Score 1654; DB 12; Length 399;
 Best Local Similarity 77.1%; Pred. No. 9.2e-154;
 Matches 300; Conservative 40; Mismatches 49; Indels 0; Gaps 0;

QY 84 KPMCAVTSPTLSPPHSLDVIISNMLMLYLSDERVENLVERMLKWLKPGGYIFFRSCF 143
 DB 10 KPMCAVTSPTLSPPHSLDVIISNMLMLYLSDERVENLVERMLKWLKPGGYIFFRSCF 69
 QY 144 HQSGDHKRSNPTHYRPFYTKAFKCHLODGSNSYELSLSCKICIGAYVRNKKONQ 203
 DB 70 HQSGDHKRSNPTHYRPFYTKAFKCHLODGSNSYELSLSCKICIGAYVRNKKONQ 129
 QY 204 ISWLMQKVDSDKDGKFORFLDTSQYKNSILRYERFVVGPGYVSTGGYTTKGFVSMDLK 263
 DB 130 ICWIMQKVSSEDRGFRFLDNVQYKSGILRYERFVVGPGYVSTGGYTTKGFVSMDLK 189
 QY 264 PQOKVLDVCGGIGGDFYMAETFDVVGFDLSVNNISPALERSIGLKCAVEFEVADCTK 323
 DB 190 PQOKVLDVCGGIGGDFYMAETFDVVGFDLSVNNISPALERSIGLKCAVEFEVADCTK 249
 QY 324 INYPNPSDVIYSRDTILHIQOKPALFRSPFVWLKPGQKVLISDYCKKAGPPSPFAAYI 383
 DB 250 KHYPNPSDVIYSRDTILHIQOKPALFRSPFVWLKPGQKVLISDYCKKAGPPSPFAAYI 309
 QY 384 KORGYDLHDVKEYGOMLKADGAFVDVLAEDRTQFIIRVLRKELETVEKEKDVIFSDPSEED 443
 DB 310 KORGYDLHDVKEYGOMLKADGAFVDVLAEDRTQFIIRVLRKELETVEKEKDVIFSDPSEED 369
 QY 444 YNDIVGWNKDLARTAKGEORWGLFVAKK 472
 DB 370 YDDIVGWNKDLARTAKGEORWGLFVAKK 398

RESULT 6
 US-10-425-114-69972
 ; Sequence 69972, Application US/10425114
 ; Publication No. US20040034888AL
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59972
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTBSINTE108803_FLI.pep
; US-10-425-114-69972

Query Match      52.5%; Score 1321; DB 12; Length 314;
Best Local Similarity 76.8%; Pred. No. 4e-121;
Matches 241; Conservative 33; Mismatches 40; Indels 0; Gaps 0;

Qy 159 REPRFTYKAFKCHODGSGNSYLSLCKCIGAYVKNKQNOISLWQKVDGK 218
Db 1 REPRFTYKAFKCHODGSGNSYLSLCKCIGAYVKNKQNOISLWQKVDGK 60
Qy 219 FORPLDTQYKCNILRYRVFGPGYVSTGGYETTKFVSMCLKPGQKLVGCGIGG 278
Db 61 FORFLDNVQYKTSGLIRYRVFGPGYVSTGGYETTKFVSMCLKPGQKLVGCGIGG 120
Qy 279 DFYMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDNGS 338
Db 121 DFYMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDNGS 180
Qy 339 TILHIQDKPALFRSPFKWLPKGGKVLISDYCKKAGPPSPFAAYIKORGYLDH 398
Db 181 TILHIQDKPALFRSPFKWLPKGGKVLISDYCKKAGPPSPFAAYIKORGYLDH 240
Qy 399 MLKADGFDVLAEDRTEQFIRVLKKELETVEKEKDVFIQSDSEEDYNDIV 458
Db 241 MLKADGFDVLAEDRTEQFIRVLKKELETVEKEKDVFIQSDSEEDYNDIV 300
Qy 459 AKGEORWGLFVAKK 472
Db 301 SAGEQRWGLFIATK 314

RESULT 7
US-10-425-114-44987
; Sequence 44987, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44987
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700211781_FLI.pep
; US-10-425-114-44987

Query Match      43.3%; Score 1088; DB 12; Length 331;
Best Local Similarity 64.8%; Pred. No. 3.8e-98;
Matches 210; Conservative 38; Mismatches 68; Indels 8; Gaps 3;

Qy 153 SNPTH-----YRPRFTYKAFKCHODGSGNSYLSLCKCIGAYVKNKQNOISLW 208
Db 12 SKPTVLLIPLDVIISNAP-ECNQRKPEGSAT---LNGLAFTKIGIKFWISVCLW 67
Qy 209 QKVDKDDKGFORPLDTSQYKCNILRYRVFGPGYVSTGGYETTKFVSMCLKPGQK 268

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59972
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTBSINTE108803_FLI.pep
; US-10-425-114-69972

Query Match      52.5%; Score 1321; DB 12; Length 314;
Best Local Similarity 76.8%; Pred. No. 4e-121;
Matches 241; Conservative 33; Mismatches 40; Indels 0; Gaps 0;

Qy 159 REPRFTYKAFKCHODGSGNSYLSLCKCIGAYVKNKQNOISLWQKVDGK 218
Db 1 REPRFTYKAFKCHODGSGNSYLSLCKCIGAYVKNKQNOISLWQKVDGK 60
Qy 219 FORPLDTQYKCNILRYRVFGPGYVSTGGYETTKFVSMCLKPGQKLVGCGIGG 278
Db 61 FORFLDNVQYKTSGLIRYRVFGPGYVSTGGYETTKFVSMCLKPGQKLVGCGIGG 120
Qy 279 DFYMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDNGS 338
Db 121 DFYMAETFDVGVGFDLSVNMISFALERSIGLKCAVEFVADCTKINYPDNGS 180
Qy 339 TILHIQDKPALFRSPFKWLPKGGKVLISDYCKKAGPPSPFAAYIKORGYLDH 398
Db 181 TILHIQDKPALFRSPFKWLPKGGKVLISDYCKKAGPPSPFAAYIKORGYLDH 240
Qy 399 MLKADGFDVLAEDRTEQFIRVLKKELETVEKEKDVFIQSDSEEDYNDIV 458
Db 241 MLKADGFDVLAEDRTEQFIRVLKKELETVEKEKDVFIQSDSEEDYNDIV 300
Qy 459 AKGEORWGLFVAKK 472
Db 301 SAGEQRWGLFIATK 314

RESULT 8
US-10-425-114-60338
; Sequence 60338, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60338
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3607-019-D6_FLI.pep
; US-10-425-114-60338

Query Match      42.3%; Score 1064; DB 12; Length 377;
Best Local Similarity 71.0%; Pred. No. 1.1e-95;
Matches 196; Conservative 33; Mismatches 47; Indels 0; Gaps 0;

Qy 1 HTVDLTIRAMLDLSOASDLKKEPEREILSMPLPGLGKLELGGAGIGRTGELAERAGOV 60
Db 80 HSGELMLEAIMLDSRAALDKKEPEREILSMPLPGLGKLELGGAGIGRTGELAERAGOV 139
Qy 61 IALDPTIESAIKKNVINGHYKXKVFMCADVTSPTLSPPHSLDVIFSNMLLYLSERVE 120
Db 140 FAVDFVESVIKKNGSINDHYGNTSPMCADVTSPTLSPPHSLDVIFSNMLLYLSERVE 199
Qy 121 NLVERMLKWLKPGGYIFFRSCFHSQSGDKKSNPHYRPRFTYKAFKCHODGSGNS 180
Db 200 KLVERNVLKLVGGYIFFRSCFHSQSGDKKSNPHYRPRFTYKAFKCHODGSGNS 259
Qy 181 YELSLSCICIGAYVKNKQNOISLWQKVDKDDKGFORPLDTSQYKCNILRYRV 240
Db 260 FKLSLITFKICIGAYVKNKQNOISLWQKVDKDDKGFORPLDTSQYKCNILRYRV 319
Qy 241 GPGYVSTGGYETTKFVSMCLKPGQKLVGCGIG 276
Db 320 GDGYVSTGGYETTKFVSMCLKPGQKLVGCGIG 355

RESULT 9
US-10-425-114-53753
; Sequence 53753, Application US/10425114
```

Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack B.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425.114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 53753
LENGTH: 219
TYPE: PRT
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3829-031-F4_FLI.pap
US-10-425-114-53753

Query Match 37.3%; Score 937; DB 12; Length 219;
Best Local Similarity 80.7%; Pred. No. 1.5e-83;
Matches 176; Conservative 16; Mismatches 26; Indels 0; Gaps 0;

QY 255 EFVSMDLKPGQKVLVDCGCGGGGDFYMAETFDVVEVGFDSLVNMISFALERSIGLKCAV 314
DB 1 EFVAKDLKPGQKVLVDCGCGGGGDFYMAKEFGYVVGIDLSINMISFALERANGIDCSV 60

QY 315 EFVADCTKINYPNSFDVIYSRDTILHIQDKPALFRSFKYKMLKPGQKVLISDYCKKAGP 374
DB 61 EFVADCTKINYPNSFDVIYSRDTILHIQDKPALFRSFKYKMLKPGQKVLISDYCKSAGA 120

QY 375 PSPEFAAYIKQGYDLHVKYQGMKADAGFVDVLABDRTEQIRVLRKLETVKKEKDV 434
DB 121 SSPEFAAYIKQGYDLHVKYQGMKADAGFVDVLABDRTEQIRVLRKLETVKKEKDV 180

QY 435 FIDSEEDYNDIVGNDKLRRTAKGEQRMGLFVAKK 472
DB 181 FIDFSEEDYNDIVGNDKLRRTAKGEQRMGLFVAKK 218

RESULT 10
US-10-424-599-257495
Sequence 257495, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424.599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 257495
LENGTH: 237
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_74541C.1.pap
US-10-424-599-257495

Query Match 33.4%; Score 839; DB 12; Length 237;
Best Local Similarity 66.7%; Pred. No. 7.7e-74;
Matches 158; Conservative 30; Mismatches 49; Indels 0; Gaps 0;

QY 237 ERVFGPGYVSTGGYETTKKFFYMLDKPGQKVLVDCGCGGGGDFYMAETFDVVEVGFDS 296
DB 1 ESVFGPGYVSTGGYETTKKFFYMLDKPGQKVLVDCGCGGGGDFYMAETFDVVEVGFDS 60

QY 297 VNMISFALERSIGLKCAVFEVADCTKINYPNSFDVIYSRDTILHIQDKPALFRSFKYK 356
DB 61 INIISLAIERAIGLKCCVFEACADCTKCTFPVNTFPIVYSRDTILHIQDKPALFRSFKYK 120

QY 357 LKPGQKVLISDYCKKAGPSPSFAAYIKQGYDLHVKYQGMKADAGFVDVLABDRTEQ 416
DB 121 LKRGGTLITDYCKSEGSLSIGYAEYIKKGYVYHDMKTCRMLNAGFDDVVAEDRTNL 180

QY 417 FIVLRKLETVKKEKDVISDFSEEDYNDIVGNDKLRRTAKGEQRMGLFVAKK 473
DB 181 FAKTLOQELNALHKSODFDIDFSEEDYNEITERWAKQKMGADGEGTWSLFIATKE 237

RESULT 11
US-10-437-963-110035
Sequence 110035, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437.963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 110035
LENGTH: 271
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(271)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_14137C.1.pap
US-10-437-963-110035

Query Match 27.5%; Score 690.5; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 3.9e-59;
Matches 134; Conservative 20; Mismatches 28; Indels 19; Gaps 2;

QY 1 HTVDLTIEANMLDSQASDLDKERPE-----ILSMLPPEEGKLEL 42
DB 71 HSKDLTVEANMLDSRAADLDKERPEFRILQSFSSPMGKLLFVLSVLPYKGSVLEL 130

QY 43 GAGIGFT-GELEAKAGVIALDFISAIKQNEVINGHYKVKFMCADVTSPILSFPFHS 101
DB 131 GAGIGFTVGTGAKKLAHVIALDFISVKNENINGHHKNITFCADVTSPDLTIEDNS 190

QY 102 LDVIFSNWLLMYLSDDEVNLMKMLKPGQYIFPESCFHQSGDHKXSNPTHREP 161
DB 191 IDLIFSNWLLMYLSDDEVNLMKMLKVGCHIFPESCFHQSGDKRKNVPTHREP 250

QY 162 RPYTKAPKECHLDGSGNSYE 182
DB 251 RPYTKIFKECHSYDKDGGPFD 271

RESULT 12
US-10-424-599-240673
Sequence 240673, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K

APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 240673
LENGTH: 183
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(183)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_59355C.1.pap
US-10-424-599-240673

Query Match 26.5%; Score 665.5; DB 12; Length 183;
Best Local Similarity 72.6%; Pred. No. 6.3e-57;
Matches 122; Conservative 14; Mismatches 25; Indels 7; Gaps 1;

Qy 104 VTFNMLLYLSDSEVENLVERMLKPKGYIPFRESCHQSGDHKRSNPTHYREPRF 163
Db 20 ITFSNMLLYLSDSEVENLVERMLKPKGYIPFRESCHQSGDHKRSNPTHYREPRF 79

Qy 164 YTKAFKECHLQDGSNGVELSLSCCKICGAYVRNKKQNIQISMLWOKVSKDDKQFQRLP 223
Db 80 YTKAFKECHLQDGSNGVELSLSCCKICGAYVRNKKQNIQISMLWOKVSKDDKQFQRLP 132

Qy 224 DTQYKCNILYRVRFGVYVSTGGYETTKFVSMLDLKPQGVLDV 271
Db 133 DRYEYSKILASERMYGPGVSTGLETTEKFPVTLGLRQGVLDV 180

RESULT 13
US-10-149-759-50
Sequence 50, Application US/10149759
Publication No. US20030157592A1
GENERAL INFORMATION:
APPLICANT: Lerchl, Jens
APPLICANT: Renz, Andreas
APPLICANT: Ehrhardt, Thomas
APPLICANT: Reindl, Andreas
APPLICANT: Clampus, Petra
TITLE OF INVENTION: Moss Genes from Physcomitrella patens encoding proteins
TITLE OF INVENTION: involved in the synthesis of tocopherols and
FILE REFERENCE: BASE/NAE 1333/99 PCT/US
CURRENT APPLICATION NUMBER: US/10/149,759
CURRENT FILING DATE: 2002-10-17
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 82
SOFTWARE: WordPerfect version 6.1
SEQ ID NO 50
LENGTH: 168
TYPE: PRT
ORGANISM: Physcomitrella patens
US-10-149-759-50

Query Match 26.2%; Score 658; DB 14; Length 168;
Best Local Similarity 72.9%; Pred. No. 3e-56;
Matches 121; Conservative 19; Mismatches 26; Indels 0; Gaps 0;

Qy 307 SIGLKCAVEFEVADCTKINYPDNPDIYSRDTLHIOKDALFESVYMLKPGGVLLIS 366
Db 1 SIARKCAVEFEVADCTKINYPDNPDIYSRDTLHIOKDALFESVYMLKPGGVLLIS 60

Qy 367 DYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGVDVLAEDRTQFIRVLKLE 426

Db 61 DYCRAPQTPSAEPAAYIQQGYDLHVSQKYGKMLBDAQFVFEVVAEDRTQFIRVLORELA 120

Qy 427 TVEKEKDVFIISDFSEYDINDVIGMNDKLRRRTAKGEQRMGLFVAKK 472
Db 121 TTEAGRDQFINDFSEYDINYVSGWKSJLRCGSDNQKGLFIAYK 166

RESULT 14
US-10-369-493-6306
Sequence 6306, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 6306
LENGTH: 437
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-6306

Query Match 20.5%; Score 516.5; DB 15; Length 437;
Best Local Similarity 33.6%; Pred. No. 1.1e-41;
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;

Qy 104 VIFSNMLLYLSDSEVENLVERMLKPKGYIPFRESCHQSGDHKRSNPTHYREPR 162
Db 68 LIFNNALSQIITNADLLTDFLNATNATAGCTVIRED-LKDCSDKQVRLTDY---- 122

Qy 163 FYTKAFKECHLQDGSNGVELSLSCCKICGAYVRNKKQNIQISMLW---QKVDSDKDKG- 218
Db 123 -----PDVFRITDSDGNNTGLDLTVDOVEH--SNVYQNFLOFIFVPRKVPAPTTDAT 175

Qy 219 --FQRLDTSQYKCNILYRVRFGVYVSTGGYETTKFVSM-LDKPGQKVLVGCQI 275
Db 176 ITRDFDLKTYTNTGIDAYEMFEGVNFISPGGYDENLKIIRFGDFKPGQTMLDIGVGI 235

Qy 276 GGGDFVMASTPDVYVVGFDLSVNMISPALESLGK-CAVEFEVADCTKINYPDNPDIYSR 334
Db 236 GGGARQVADFEVYVVGFDLSVNMISPALESLGK-CAVEFEVADCTKINYPDNPDIYSR 295

Qy 335 YSRDTLHIOKDALFESVYMLKPGGVLLISDYCKKAGPPSPFAAYIKQGYDLHDV 394
Db 296 FSRDCTQHIDPTEKLFISRIYKALPGGVLLITMYGKGYSQDKFKTYVQAQYFLK 355

Qy 395 EYQMLKADGVDVLAEDRTQFIRVLKLETVKEKDVFIISDFSEYDINDVIGMNDK 454
Db 356 EIADIANKTGFVNVQNTNTPFRFKEILAEERHLEQNEABFMSKFTQREDSLSICWTOK 415

Qy 455 LRRTAKGEQRMGLFVAKK 472
Db 416 LGYIEKDNENWAFFLAQK 433

RESULT 15
US-10-602-268-21
Sequence 21, Application US/10602268
Publication No. US20040091467A1
GENERAL INFORMATION:
APPLICANT: Williams, Deyck J.
APPLICANT: McIaird, Werty B.
APPLICANT: Hresko, Michelle Coutu
APPLICANT: Frevert, Anita M.


```
; APPLICANT: Worthington, Ronald E.
; APPLICANT: Kioek, Andrew P.
; APPLICANT: Davila-Aponte, Jennifer A.
; APPLICANT: Bradley, John D.
; APPLICANT: Xu, Siqun
; TITLE OF INVENTION: NEMATODE PHOSPHOTRANSLAMINE
; FILE REFERENCE: 12557-011001
; CURRENT APPLICATION NUMBER: US/10/602,268
; PRIOR FILING DATE: 2003-06-23
; PRIOR APPLICATION NUMBER: US 60/390,490
; FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-602-268-21

Query Match      20.5%; Score 516.5; DB 16; Length 437;
Best Local Similarity 33.6%; Pred. No. 11e-41;
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;

Qy 104 VPSNMLMYLSDEE-VENLVERMLKWLKPGGYIPFRESCTHOSGDHKKSNPTHYREPR 162
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 68 LIPNNALSQIITNADLLTDFLKNATNATAIGGTVIRED-LKDCSDKQVRLTDY---- 122
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 163 FYTKAFKECHLDQSGNSYELSLSCKICIGAYVRNKKNNQISLW---QKVDSDKDKG- 218
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 123 -----FDVPRITDSDGNTGLDLYTDOVEH--SNVVEQNFLDFIPVFRKKVFAPTTDAT 175
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 219 --FORFLDTSQYKCNILRYRVPFGYVSTGGYETTKFVSM-LDKGOKVLDVGGGI 275
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 176 ITRFDLTKQTNTGIDAYEMFVGVPISPGGYDENLKIIRFGDFKPGQMLDIGVI 235
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 276 GGGDFYMAETFDVYVVGFDLSVNMISFALERSIGLK-CAVEPEVADCTKINYPDNSFDVI 334
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 236 GGGARQVADFQGVHVGIDLSNNMLATALERLHEEKDSRVKYSITDALVYQFEDNSFDVY 295
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 335 YSRDTTLHIQDPALFRSPYKWLPGGKVLISDYCKKAGPPSPFAAYIKORGYDLHDVK 394
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 296 PSRDCIHIPTDEKLFPSRIYKALFEGKVLITMGKGYGQSDKFKTYVAQRAYFLKNLX 355
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 395 BYGQMLKADAGFDVLAEDRTQPIRVLKELETVEKEKQVPIGDPSEEDYNDIVGWNDK 454
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 356 EIADIANKTGPNVYQTNMTFRFKFILLERGHLEQNEAFMSKFTQREDRSLISGWDK 415
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 455 LRETAKGEQORGLFVAKK 472
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 416 LGVIEKDNHNNWFFLAQK 433
Db :||: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: July 26, 2004, 13:21:07
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:12:11 ; Search time 17 Seconds
(without alignments)
2676.388 Million cell updates/sec

Title: US-10-031-331b-40
Perfect score: 2515
Sequence: 1 HTVDLTIEAMLDQASDL.....KLRTAKGQRWGLFVAKKK 473
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2041	81.2	555	2 H96762	hypothetical prote
2	1294.5	51.5	374	2 F96525	protein T1N15.23 [
3	516.5	20.5	437	2 T23330	hypothetical prote
4	413	16.4	495	2 T27936	hypothetical prote
5	217.5	8.6	363	2 T06795	probable sterol 24
6	214	8.5	280	2 A22031	gamma-tocopherol m
7	210.5	8.4	317	2 S75618	hypothetical prote
8	199	7.9	344	2 T01572	sterol 24-C-methyl
9	198	7.9	344	2 T04138	sterol 24-C-methyl
10	193	7.7	367	2 T06780	probable sterol 24
11	187.5	7.5	330	2 AC2071	hypothetical prote
12	185.5	7.4	227	2 C71246	hypothetical prote
13	184.5	7.3	318	2 S76226	hypothetical prote
14	183.5	7.3	387	2 A81324	cyclopropane-fatty
15	180.5	7.2	306	2 S18533	eryg protein - Sac
16	179	7.1	346	2 T10173	sterol 24-C-methyl
17	177.5	7.1	462	2 B75305	probable cycloprop
18	175	7.0	283	2 T44579	CS-O-methyltransferase
19	173.5	6.9	227	2 C75210	sterol biosynthesis
20	172.5	6.9	348	2 C96673	gamma-tocopherol m
21	171.5	6.8	303	2 A2776	methyltransferase
22	171.5	6.8	321	2 C37556	probable methyltra
23	171	6.8	270	2 A11873	hypothetical prote
24	170	6.8	283	2 J6531	avermectin B 5-O-m
25	169.5	6.7	244	2 A2704	methyltransferase
26	169.5	6.7	244	2 H37485	biotin synthetase p
27	168	6.7	379	2 T50969	probable DELTA(24)
28	167.5	6.7	241	2 G89925	menaquinone biosyn
29	167	6.6	218	2 H84364	membrane protein [

ALIGNMENTS

RESULT 1

H96762

hypothetical protein FeD5.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: H96762

R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.;

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.R.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Mafti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: H96762

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-555 -CSTO-

A:Cross-references: GB:AB005173; NID:g10092368; PIDN:AGI2776.1; GSPDB:GM00141

C:Genetics:

A:Gene: FeD5.1

A:Map position: 1

Query Match 81.2%; Score 2041; DB 2; Length 555;

Best Local Similarity 78.4%; Pred. No. 6.7e-145;

Matches 370; Conservative 53; Mismatches 49; Indels 0; Gaps 0;

QY 1 HTVDLTIEAMLDQASDLKKEPELISMLPLEGKCLLELGNGIGRTGLAKAGQV 60

DB 83 HVGVSUVAAMLDKASDLKKEPELIFAPPIGTITVLEFGAGIGRTGLAKAGQV 142

QY 61 IALDFIRSAIKKNEVINGHYKNVKNPCADVTSTPLTFPPHSLDVIFSNWLLMYLSDREEVE 120

DB 143 IAVDFIESVIKKNEVINGHYKNVKNPCADVTSPNNMPPNEMDLIFSNWLLMYLSDQVEE 202

QY 121 NVEVRLMKLPGGYIFPSCFHOSDHRKKNPHYRPRPYTKAPKECHLQDSGNS 180

DB 203 DLAKNQLQWTKVGYIFPSCFHOSDHRKKNPHYRPRPYTKAPKECHLQDSGNS 262

QY 181 YELSLKSCICGAVYRNKNQNSIWLKQVSKDKGQFORFLDTSQYKNSILYERVF 240

DB 263 YELSLVSKCICGAVYRNKNQNSIWLKQVSKDKGQFORFLDTSQYKNSILYERVF 322

QY 241 GGYVSTGGYETTKFVSMLDLKGQKVLVGGIGGDPFYMAETFDVWVGFDLSVNNI 300

DB 323 GGFVSTGGLETKEFVMDLKGQKVLVGGIGGDPFYMAETFDVWVGFDLSVNNI 382

QY 301 SFALRSRGLKCAVEFEVADCTKINYPDNSFDVIYSDRTILHIDQKDALFRSFKWLKFG 360

conserved hypothet
probable sterol 24
sterol 24-C-methyl
ubiquinone/menaqui
hypothetical prote
cyclopropane-fatty
cyclopropane fatty
methyltransferase
hypothetical prote
cyclopropane-fatty
conserved hypothet
probable methyltra
probable sterol 24
sterol methyltrans

30 167 6.6 391 2 T08338
31 165.5 6.6 352 2 T03845
32 165.5 6.6 357 2 T03848
33 163.5 6.5 361 2 S63686
34 161.5 6.4 260 2 D82366
35 161.5 6.4 366 2 T33885
36 161.5 6.4 382 2 A44292
37 161.5 6.4 382 2 B90925
38 161.5 6.4 382 2 F85773
39 160.5 6.4 274 2 AC2973
40 160.5 6.4 274 2 D83309
41 160 6.4 389 2 H64571
42 160 6.4 542 2 AB1066
43 159.5 6.3 269 2 C75569
44 159.5 6.3 324 2 T42375
45 159.5 6.3 378 2 T39579

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29330
R:Pauley, A.; Gattung, S.
A:Description: The sequence of C. elegans cosmid F54D11.
A:Reference number: 220606
A:Accession: T29330
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-437 <PAU>
A:Cross-references: EMBL:U64834; PIDN:AB04824.1; GSPDB:GN00023; CESP:F54D11.1
A:Experimental source: strain Bristol N2; clone F54D11
C:Genetics:
A:Gene: CESP:F54D11.1
A:Map position: 5
A:Introns: 36/2; 76/3; 108/1; 315/3; 398/3
Query Match 20.5%; Score 516.5; DB 2; Length 437;
Best Local Similarity 33.6%; Pred. No. 6.5e-31;
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;
QY 104 VIFSNLLMYLSDEE-VENLVVERMLKPGGKGVVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 394
DB 68 LIFNNALSQIITNADLLTDFLKNATNATAGTGTIRED-LKDCSKRQVRLTDT- 122
QY 163 FYTKAPKECHLDQSGNSYELSLSCIKIGAYVRNKKQNIQISLW- --QKVDKDDKG- 218
DB 123 -----FDVPRITDSDGNNLTGLDVTVDVEH--SNTVEQNFLDFIVFRKCVFAPTDDAT 175
QY 219 --PQRLDTSQYKCNILRYERFVPGVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 275
DB 176 ITRDFELDKTQVNTGIDAYEMFVGFISPGYDENLKIIFGDFKFGQTMLDIGVGI 235
QY 276 GGGDFYMAETPDVGVGFDLSVNMISFALERSIGLK-CAVEPEVADCTKINYDNPSFDVI 334
DB 236 GGGARQVADFVGVHVGIDLSNNMLATLALERHBEKDSRVKYSITDALVTQFEDNSFDIV 295
QY 335 YSRDTLHTQDPALFRSPYKMLKPGGKGVVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 394
DB 296 FSRDCHQHPDTEKLFPSRIYKALKPGGKVLITWYKGVGEOSDKFTYVAGRAYFLANLK 355
QY 395 EYQMLKADAGFDVLAEDTEQFIATKELTETVEKEDVPISEEDVNDIVGVWMDK 454
DB 356 EIADIANKTQFVNVOTENNTPRFKEILLERHCHLSEAEAPSKFTQERDLSLISGWTDK 415
QY 455 LRRTAKGEORWGLFVAKK 472
DB 416 LGVIEKDNHNNWFFLAQK 433

RESULT 4
T29336
hypothetical protein ZK622.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T29336
R:Leimbach, D.
A:Description: The sequence of C. elegans cosmid ZK622.
A:Reference number: 220443
A:Accession: T29336
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-495 <LEI>
A:Cross-references: EMBL:U99998; PIDN:AAA81102.1; CESP:ZK622.3
C:Genetics:
A:Gene: CESP:ZK622.3
A:Introns: 34/2; 97/2; 222/3; 310/3
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK622.3
Query Match 16.4%; Score 413; DB 2; Length 495;
Best Local Similarity 28.6%; Pred. No. 4.3e-23;

393 SFALHAIGLKCSVEFEVADCTKKEYPDNTFDVYISRTILHIQDKPALFRFFYKWLKPG 442
QY 361 GKVLISYCKKAGPSPSPAAITKORGYLHIVKEYGQMLKADAGFDVLAEDTEQIRV 420
DB 443 GKVLITDYSRPTKPSDFPFAITKORGYLHIVKEYGQMLKADAGFEVLAEDTRIDQFMKV 502
QY 421 LRKELETVKESKQVIFSDPSEEDVNDIVGWNDDKLRRTAKGEORWGLFVAKK 472
DB 503 LKRELDAVEKEKEFISDFSKEDYEDIGGWSKILRSGSGQKWLGTAKR 554
RESULT 2
F96525
protein T1N15.23 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96525
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizcar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, X.; Liu, X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzitali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.K.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <STO>
A:Cross-references: GB:AB005173; NID:98778697; PIDN:AAF97905.1; GSPDB:GN00141
C:Genetics:
A:Gene: T1N15.23
A:Map position: 1
Query Match 51.5%; Score 1294.5; DB 2; Length 374;
Best Local Similarity 66.8%; Pred. No. 3e-89;
Matches 245; Conservative 30; Mismatches 33; Indels 59; Gaps 2;
QY 166 KAPKECHLDQSGNSYELSLSCIKIGAYVRNKKQNIQISLW- --QKVDKDDKG- 225
DB 8 QVQECQCDASGNSFELSMVCKCIGAYVRNKKQNIQISLW- --QKVDKDDKG- 67
QY 226 SQYKCNILRYERFVPGVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 275
DB 68 VQYKSGILRYERFVPGVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 235
QY 256 FVSMMLDKPQKQVLDVGVGIGGDFYMAETPDVGVGFDLSVNMISFALERSIGLKCAVE 315
DB 128 FVAMWDLKPGQKQVLDVGVGIGGDFYMAETPDVGVGFDLSVNMISFALERSIGLKCAVE 187
QY 316 FEVADCTKINYDNPSFDVYISRTILHIQDKPALFRSPYKMLKPGGKGVVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 375
DB 188 FEVADCTKINYDNPSFDVYISRTILHIQDKPALFRSPYKMLKPGGKGVVSTGGVETTKFVSM-L-DLKPQKQVLDVGVK 247
QY 376 SPFAAYIKORGYLHIVKEYGQMLKADAGFDVLAEDTEQIRVLRKELETVKESKQV 435
DB 248 SPFAAYIKORGYLHIVKEYGQMLKADAGFDVLAEDTEQIRVLRKELETVKESKQV 307
QY 436 ISDFSE-----EDVNDIVGVWMDKLRRTAKGEORW 466
DB 308 ISDFSEVKTTLTNSFQNTCLLCHDHLKLLQDNDIVGVWMDKLRRTAKGEORW 367
QY 467 LFVAKK 473
DB 368 LFVAKK 374

RESULT 3
T29330
hypothetical protein F54D11.1 - Caenorhabditis elegans

Db 49 WEDVGHMHGTYGFGYRDRQAQIDLIKELLANAVPQNSAKP-RKILDLGGGIGG 107
 QY 278 GDFMAETFDVEVGGDLSVNMISPALE--RSIGLKCAVEFVADCTKINYPNSPDVY 335
 Db 108 SLLYLAQOHAQVNGASLSVQVERAGRARALGIGSTCOFQVANALDLPASDFDWW 167
 QY 336 SRDILHIQKPALEFRFYKWLKPGGKVLISDYCKA-----GPPSPFAAYIKQRYDL 390
 Db 168 SLESGEHMKNAQLOEAWNLKPGGRLILATWCHRPIDPQNGPLTADERRHL-QAIYDV 226
 QY 391 H-----DVKEYGQMLKADGPDVLAED 412
 Db 227 YCLFVVSPLPYEAIARECGFGKTKAD 254

RESULT 8
 T01572
 sterol 24-C-methyltransferase (EC 2.1.1.41) - maize
 N:Alternate names: (S)-adenosyl-L-methionine:delta 24-sterol methyltransferase
 C:Species: Zea mays (maize)
 C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 03-Jun-2002
 C:Accession: T01572
 R:Tseng, Y.; Nes, W.D.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: Z14350
 A:Accession: T01572
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-344 <TON>
 A:Cross-references: EMBL:AF045570; NID:G2909845; PIDN:AAC04265.1; PID:G2909846
 A:Experimental source: strain B73
 C:Genetics:
 A:Gene: SMT
 C:Superfamily: 24-sterol C-methyltransferase; bioC homology
 C:Keywords: methyltransferase; S-adenosylmethionine
 F:101-205/Domain: bioC homology <BIOC>

Query Match 7.9%; Score 199; DB 2; Length 344;
 Best Local Similarity 34.5%; Pred. No. 2.8e-07;
 Matches 60; Conservative 26; Mismatches 72; Indels 16; Gaps 7;
 QY 254 KEFVSM-LDLKPGQKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGL 310
 Db 89 EHFLALQLGLKPGKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGL 148
 QY 311 KCAVEFVADCTKINYPNSPDVYSRDILHIQKPALEFRFYKWLKPGGKVLISDYCK 370
 Db 149 SGTCDVFVKADEFMPPDDNTFDVAYIAETCHAPDPVGCYKEIYRVLKPGQCFAYVWC- 207
 QY 371 KAGPPSPFAAY--IK---ORGYDLHDVKEYGQML---KDAGFVDV---LAED 412
 Db 208 VTDHYDPNNAHTRIKDBIELGNGLPDIRSTRQCLRAVKADGAFVVDKDLAED 261

Query Match 7.9%; Score 199; DB 2; Length 344;
 Best Local Similarity 34.5%; Pred. No. 2.8e-07;
 Matches 60; Conservative 26; Mismatches 72; Indels 16; Gaps 7;
 QY 254 KEFVSM-LDLKPGQKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGL 310
 Db 89 EHFLALQLGLKPGKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGL 148
 QY 311 KCAVEFVADCTKINYPNSPDVYSRDILHIQKPALEFRFYKWLKPGGKVLISDYCK 370
 Db 149 SGTCDVFVKADEFMPPDDNTFDVAYIAETCHAPDPVGCYKEIYRVLKPGQCFAYVWC- 207
 QY 371 KAGPPSPFAAY--IK---ORGYDLHDVKEYGQML---KDAGFVDV---LAED 412
 Db 208 VTDHYDPNNAHTRIKDBIELGNGLPDIRSTRQCLRAVKADGAFVVDKDLAED 261

RESULT 9
 T04138
 sterol 24-C-methyltransferase (EC 2.1.1.41) ESM1, endosperm - maize
 N:Alternate names: C-24 sterol methyltransferase
 C:Species: Zea mays (maize)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
 C:Accession: T04138
 R:Grebenok, R.J.; Galbraith, D.W.; Dellapenna, D.
 Plant Mol. Biol. 34: 891-896, 1997
 A:Title: Characterization of zea-mays endosperm C-24 sterol methyltransferase - one of 2
 A:Reference number: Z09668; MUID:97435974; PMID:9290641
 A:Accession: T04138
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-344 <GRE>
 A:Cross-references: EMBL:U79669; NID:G1899059; PIDN:AAB70886.1; PID:G1899060
 A:Experimental source: endosperm
 C:Genetics:

A:Gene: ESM1
 C:Function:
 A:Description: methyltransferase
 C:Superfamily: 24-sterol C-methyltransferase; bioC homology
 C:Keywords: methyltransferase; S-adenosylmethionine
 F:101-205/Domain: bioC homology <BIOC>
 Query Match 7.9%; Score 198; DB 2; Length 344;
 Best Local Similarity 34.4%; Pred. No. 3.3e-07;
 Matches 59; Conservative 26; Mismatches 74; Indels 14; Gaps 6;
 QY 254 KEFVSM-LDLKPGQKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGL 310
 Db 89 EHFLALQLGLKPGKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGL 148
 QY 311 KCAVEFVADCTKINYPNSPDVYSRDILHIQKPALEFRFYKWLKPGGKVLISDYCK- 369
 Db 149 SGTCDVFVKADEFMPPDDNTFDVAYIAETCHAPDPVGCYKEIYRVLKPGQCFAYVWC- 208
 QY 370 KAGPPSPFAAYIK---ORGYDLHDVKEYGQML---KDAGFVDV---LAED 412
 Db 209 TDDHYDPNNAHTRIKDBIELGNGLPDIRSTRQCLRAVKADGAFVVDKDLAED 261

RESULT 10
 T06780
 probable sterol 24-C-methyltransferase (EC 2.1.1.41) - soybean
 N:Alternate names: S-adenosyl-L-methionine:delta24-sterol-C-methyltransferase
 C:Species: Glycine max (soybean)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
 C:Accession: T06780
 R:Shi, J.; Gonzales, R.A.; Bhattacharyya, M.K.
 J. Biol. Chem. 271, 9384-9389, 1996
 A:Title: Identification and characterization of an S-adenosyl-L-methionine: delta 24-ster
 A:Reference number: Z15807; MUID:96199190; PMID:8621604
 A:Accession: T06780
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <SHI>
 A:Cross-references: EMBL:U43683; NID:G1399379; PIDN:AAB04057.1; PID:G1399380
 A:Experimental source: Cultivar Williams 82; etiolated hypocotyls
 C:Superfamily: 24-sterol C-methyltransferase; bioC homology
 C:Keywords: methyltransferase; S-adenosylmethionine
 F:122-226/Domain: bioC homology <BIOC>

Query Match 7.7%; Score 193; DB 2; Length 367;
 Best Local Similarity 27.2%; Pred. No. 8.5e-07;
 Matches 63; Conservative 35; Mismatches 90; Indels 44; Gaps 7;
 QY 260 LDLKPGQKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGLKCAVEFE 317
 Db 117 LGLKPGQKVLVDGCGGDFYMAETFDVEVGGDLSVNMISPA--LERSIGLKCAVEFE 176
 QY 318 VADCTKINYPNSPDVYSRDILHIQKPALEFRFYKWLKPGGKVLISDYCK--KAGPP 375
 Db 177 KADFMKMPDPSFDVAYIAETCHAPDPVGCYKEIYRVLKPGQCFAYVWCMTDSFDQ 236
 QY 376 SPEFAAYIK---ORGYDLHDVCK---EYGOMLKADGAFVVDVLAEDR----- 413
 Db 237 NPEHOK-IAKEIEIGDGLPDIRLTAKLEALKQAGFEVWEKDLAVDSPLPWLPLDKSH 295
 QY 414 -----TEQPIRVLRKELETVEKEDVPSIDPSREDYNDIVGG 450
 Db 296 FSLSSFLTAVALGRLFTKNMKVL--EYVGLAPKGLSLAVQDFLEKAGGLVEG 345

RESULT 11
 AC2071
 hypothetical protein all2121 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC-7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AC2071

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchida, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001.
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain A1807; MUID:21595285; PMID:11759840
A:Reference number: AC2071
A:Accession: S76226
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA073820.1; PID:g17131212; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2121

Query Match 7.5%; Score 187.5; DB 2; Length 330;
Best Local Similarity 29.1%; Pred. No. 1.9e-06;
Matches 48; Conservative 35; Mismatches 71; Indels 11; Gaps 4;
QY 262 LKPGQKVLVCGGIGGDFYMAETFDVVGFDLSVNMISFALERSIGLKCAVEFEVADC 321
DB 90 LPPGTTLLDVCGGSSRILARDYGFVAVTGITISPOQVQRAQELT-PQELNAQFLVDDA 148
QY 322 TKINYPDMSFDVYISRDTHLHODKPAIPRSFYKMLKPGKVLISDYCKKAGPSP--EF 379
DB 149 MALSPFNSFDVNSIEAGPMPDPAIPAKELMRVLKPGGIVLADWQDRDQKPLNF 208
QY 380 AAYIKQGYD-----LHDVKEYGQMLKQAGFY--DVLARDRTQ 416
DB 209 EKPVNQQLLDQWHPAFSSIEGFSSELLAATGLVEGEVITADWTQ 253

RESULT 12
C71246

hypothetical protein PH0226 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: C71246
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekinaka, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kuchida, N.; Oguchi, M. DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon Pyrococcus horikoshii; MUID:98344137; PMID:9679194
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: C71246
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-227 <KAM>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BA029298.1; PID:g3256615
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0226
A:Superfamily: conserved hypothetical protein MJ1123; bioC homology
F:39-144/Domain: bioC homology <BIOC>

Query Match 7.4%; Score 185.5; DB 2; Length 227;
Best Local Similarity 35.8%; Pred. No. 1.6e-06;
Matches 49; Conservative 24; Mismatches 55; Indels 9; Gaps 4;
QY 236 YRVFPGYVSTGYE-----TTKEFVSMLDLKPQKVLVCGGIGGDFYMAETFDVEV 290
DB 6 YRVF-PYTYDINSQEVRSRTTLEPLLMKMKKGVLDLACGVGFSF-LLEDYGEV 63
QY 291 VGFDSVNMISFALERSIGLKCAVEFEVADCTKINYPDMSFDVYISRDTHLHIO--DKPA 348
DB 64 VGVDSIEDMKAREYAKRSNVEFVVGARKLSFEDKTFDVIPIFDSIVHFEPLNQ 123
QY 349 LFRSFYKMLKPGKVL 365
DB 124 VFKEVRRVLKPSGKFM 140

RESULT 13
S76226

hypothetical protein - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M. DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp. strain PCC 6803
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76226
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-318 <KAN>
A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BA018485.1; PID:d101921
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: bioC homology
F:94-195/Domain: bioC homology <BIOC>

Query Match 7.3%; Score 184.5; DB 2; Length 318;
Best Local Similarity 29.6%; Pred. No. 3e-06;
Matches 47; Conservative 30; Mismatches 61; Indels 21; Gaps 5;
QY 227 QYKNSILRYRVERFGYVSTGY---ETTKFV-SMLD-----LKPQKVL 269
DB 42 QWTEGILEY---YMGDHIHLGHYDPPVAKDFIOSKIDFVHMAQWGLDLPPTVL 98
QY 270 DVGGIGGDFYMAETFDVVGFDLSVNMISFALERSIGLKCAVEFEVADCTKINYPDN 329
DB 99 DVGGIGGSSRILAKDYGFNVGTITISPOQVQRAQELT-PEDVTAKFAVDAMALSPDG 157
QY 330 SFDVYISRDTHLHODKPAIPRSFYKMLKPGKVLISDY 368
DB 158 SFDVYISRDTHLHODKPAIPRSFYKMLKPGKVLISDY 368

RESULT 14
A81324

cyclopropane-fatty-acyl-phospholipid synthase (EC 2.1.1.79) Cj1183c [similarity] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: A81324
R:Parhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin, C.W.; Quail, M.; Rajadream, M.A.; Rutherford, K.M.; VanVleet, A.; Whithead, S.; Barrel, N. Nature 403, 663-669, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: A81324
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-387 <PAR>
A:Cross-references: GB:AL139077; GB:AL111169; NID:g6968444; PIDN:CA073437.1; PID:g696861
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: cfa; Cj1183c
C:Superfamily: cyclopropane-fatty-acyl-phospholipid synthase; MCM homology
C:Keywords: methyltransferase; phospholipid metabolism; S-adenosylmethionine

Query Match 7.3%; Score 183.5; DB 2; Length 387;
Best Local Similarity 25.2%; Pred. No. 4.7e-06;
Matches 74; Conservative 51; Mismatches 114; Indels 55; Gaps 13;
QY 198 KKNQNSLWLVKQVSDKDG-----FQFL-DTSQYK-----NSILRYRVERFGP 242
DB 91 KTKQDLVSKITQKESQNIKSHYDIGNDFYKWLDDTMSYCAYPKFNNTL-YE----- 144
QY 243 GYVSTGYETTKFVSMLDLKPQKVLVCGGIGGDFYMAETFDVVGFDLSVNMISF 302
DB 145-----AQINKTEHTLTKDLDKEGKLDIGCGWGLSTMAAQYKGVVGVGITISSEQCK 199
QY 303 ALER--SIGLKCAVEFEVADCTKINYPDMSFDVYISRDTHLHIO--ODKPAIFRSFYKMLK 358

Db 200 AKERVKELGLEDKIEIQLNQDLRF-ENYFDKVVSVGMFEHVGKKNLGLVFMKVQVLK 258
 QY 359 PGGKVL-SDYCKKAGPPPEFAAYIKQGYDLHDKYQOMLKDAGFVDVLAEDRTEOF 417
 Db 259 PGGKVLHSLAMPEGKTNADIKYIPGGY-LPSLRVVSAMSEWDPHLLASLRMHY 317
 QY 418 IRVLKSLVETVEKOVFISDFSESDYDNIIVGMDNKLRTAKGE--QRWGLFV 469
 Db 318 AKTL-----DLMDENFNKVL-----DKVREKYDEEFIRMWOLYL 351

RESULT 15
 S18533
 cryG protein - Saccharopolyspora erythraea
 C;Species: Saccharopolyspora erythraea
 C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 22-Oct-1999
 C;Accession: S18533; S16747
 R;Haydock, S.F.; Dowson, J.A.; Dhillon, N.; Roberts, G.A.; Cortes, J.; Leadlay, P.P.
 Mol. Gen. Genet. 230, 120-128, 1991
 A;Title: Cloning and sequence analysis of genes involved in erythromycin biosynthesis in
 methyltransferases.
 A;Reference number: S18530; MUID:92079886; PMID:1840640
 A;Accession: S18533
 A;Molecule type: DNA
 A;Residues: 1-306 <HAY>
 A;Cross-references: EMEL:X60379; NID:948941; PIDN:CAA42929.1; PID:G581650
 C;Genetics:
 A;Gene: cryG
 A;Start codon: GTG
 C;Superfamily: bioC homology
 F;82-187/Domain: bioC homology <BIOC>

Query Match 7.2%; Score 180.5; DB 2; Length 306;
 Best Local Similarity 24.5%; Pred. No. 5.8e-06;
 Matches 64; Conservative 45; Mismatches 95; Indels 57; Gaps 11;

QY 206 W-LWQKV--DSKDDKGFQRLDTSQYKCNILRYRVFGPGYVSTGGY-----ETTK 254
 Db 18 WHVWTRVPSRRALAYALFADD-----HEATTEGAYINLGTWPGCAGLEHANG 67

QY 255 EFVSM-----DLKPGQKVLDCGIGGDFYMAET-FDVEVVGFDLSVNMISFALERS-- 307
 Db 68 ELANQLAERAGISGEDEVLDVGFUGAQDFPWLSTRKPARIVGVDLTPSHVRIASERAER 127

QY 308 IGLKCAVEBEVADCTKINYPDNPDIYSDTILHIOKDPALFRSFYKMLKPGKVLISD 367
 Db 128 ENVQDRLLQFAGSADLPFGAETFDRTVTSLESALRYEPTDFFKGAFEVLKPGVLAIGD 187

QY 368 YCK-----KAGPPSPPEAAIKQGYDLH-----DVKEYGQMLKDGAVDY--- 408
 Db 188 IIPDLRERPGSDPPK-----LAPQSGSLSGGIPVENWVPRETYAKQLEBAGFVDVVK 242

QY 409 -----LAEDRTEQFINVLKKE 424
 Db 243 SVRDNWEPWLDYVLKQLQDE 263

Search completed: July 26, 2004, 13:15:54
 Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:11:16 ; Search time 13 Seconds
(without alignments)
1894.553 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2315

Sequence: 1 HTVDLTFENMLDSQSLD.....KURRTAKGRQWGLFVAKKK 473

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2204	87.6	494	1	PEAM SPIOL
2	2077	82.6	475	1	PMW2 ARATH
3	2041	81.2	490	1	PMW3 ARATH
4	2011	80.0	491	1	PMW1 ARATH
5	176.5	7.0	251	1	UBIE SHON
6	173.5	6.9	348	1	GROM ARATH
7	167.5	6.7	241	1	UBIE STAM
8	161.5	6.4	260	1	UBIE VIRCH
9	161.5	6.4	381	1	CPA ECOLI
10	160.5	6.4	251	1	UBIE ECOLI
11	160.5	6.4	259	1	UBIE VIBPA
12	159.5	6.3	378	1	ERG6 SCHPO
13	158.5	6.3	251	1	UBIE ECOLI
14	158.5	6.3	260	1	UBIE VIBVU
15	157	6.2	245	1	UBIE NEIMA
16	156.5	6.2	241	1	UBIE STREP
17	156	6.2	245	1	UBIE NEIMB
18	155.5	6.2	237	1	UBIE LISMO
19	155.5	6.2	238	1	UBIE OCEIH
20	154.5	6.1	244	1	YXBB BACSU
21	154.5	6.1	256	1	UBIE PSBP
22	154.5	6.1	376	1	ERG6 CANAL
23	153.5	6.1	253	1	UBIE XANCP
24	153	6.1	382	1	ERG6 YEAST
25	151.5	6.0	251	1	UBIE SALTY
26	151.5	6.0	251	1	UBIE YERPE
27	149.5	5.9	237	1	UBIE BACNA
28	149.5	5.9	268	1	UBIE PASMU
29	148.5	5.9	253	1	UBIE XANAC
30	148.5	5.9	256	1	UBIE PSESM
31	147.5	5.9	237	1	UBIE BACCR
32	147.5	5.9	256	1	UBIE PSEPU
33	146.5	5.8	250	1	UBIE CONSU

RESULT 1

PEAM SPIOL

ID PEAM SPIOL STANDARD; PRT; 494 AA.

AC Q9M571;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoethanolamine N-methyltransferase (EC 2.1.1.103).
GN PEAMT.

OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;

[1]

SEQUENCE FROM N.A., AND CHARACTERIZATION.

STRAIN=sv. Savoy hybrid 612; TISSUE=Leaf;

RX MEDLINE=20261526; PubMed=10799484;

RA Nuccio M.L., Ziemak M.J., Henry S.A., Weretlynyk E.A., Hanson A.D.;

RT "Phosphoethanolamine N-methyltransferase from spinach: cDNA cloning by

RT complementation in *Schizosaccharomyces pombe* and characterization of

RT the recombinant enzyme.";

RL J. Biol. Chem. 275:14095-14101(2000).

CC -I- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,

CC phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the

CC three methylation steps required to convert phosphoethanolamine to

CC phosphocholine. Mediates a key step in the biosynthesis of

CC choline, a precursor of the osmoprotectant glycine betaine. Has an

CC optimal pH of 7.8 to 8.5. Has no ethanolamine- or

CC phosphatidylethanolamine-N-methyltransferase activity.

CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine

CC phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine

CC phosphate.

CC -I- ENZYME REGULATION: Inhibited by phosphocholine but not by choline,

CC glycine betaine, monomethylethanolamine or dimethylethanolamine.

CC -I- SUBUNIT: Monomer.

CC -I- SUBCELLULAR LOCATION: Cytoplasmic.

CC -I- INDUCTION: By salt stress.

CC -I- SIMILARITY: Belongs to the methyltransferase superfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----AF237633; AAF61950.1; --

DR InterPro; IPR001601; Methyltransf.

DR InterPro; IPR000051; SAM bind.

DR Methyltransferase; Transferase; Repeat.

KW DOMAIN: 53.....I62..... SAM-BINDING 1..

FT DOMAIN 282 389 SAM-BINDING 2.

FT SEQUENCE 494 AA; 56361 MW; 7F2537C8E4B8413B CRC64;

34 146 5.8 231 1 UBIE STRCO
35 145.5 5.8 259 1 UBIE HAEDU
36 143.5 5.7 224 1 UBIE BACST
37 143.5 5.7 256 1 UBIE_PSEAS
38 141 5.6 345 1 YI37_STRFR
39 140.5 5.6 286 1 MMAL_MTCU
40 140 5.6 244 1 UBIE NITEU
41 140 5.6 246 1 UBIE MICLU
42 138 5.5 239 1 UBIG XANAC
43 136.5 5.4 230 1 UBIE CORBF
44 135.5 5.4 252 1 UBIE LACLA
45 134 5.3 231 1 UBIE_STRAW

ALIGNMENTS

Query Match 87.6%; Score 2204; DB 1; Length 494;
 Best Local Similarity 86.4%; Pred. No. 8.9e-159;
 Matches 408; Conservative 34; Mismatches 30; Indels 0; Gaps 0;

QY 1 HTVDLTTTAAWMLDSQASDLKBERPHTLSMLPPELGKCLLELGAGIGRTGELAAGQOV 60
 DB 22 HSDVDTVEAWMLDSQASDLKBERPHTLSMLPPELGKCLLELGAGIGRTGELAAGQOV 81
 QY 61 IALDPIESAIKNEVINGHYKVKFPCADVTSTLSPFPHSDLVFNSWLLMYLSDEVE 120
 DB 82 IALDPIESAIKNEVINGHYKVKFPCADVTSTLSPFPHSDLVFNSWLLMYLSDEVE 141
 QY 121 NLVERMLKWLKPGGYIFFRESCHQSDHKKSNPTHYREPRFTYKAPKCECHLQDGSNS 180
 DB 142 LRVEMMLKWLKPGGYIFFRESCHQSDHKKSNPTHYREPRFTYKAPKCECHLQDGSNS 201
 QY 181 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSQYKCNLSLYRYVP 240
 DB 202 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSQYKCNLSLYRYVP 261
 QY 241 GFGYVSTGGYETTFKVFMSMLDLKPGQKVLVGGCGIGGDFYMAETDVEVGFDSVNM 300
 DB 262 GFGYVSTGGYETTFKVFMSMLDLKPGQKVLVGGCGIGGDFYMAETDVEVGFDSVNM 321
 QY 301 SPALERSIGLCAVEFEVADCTKINYPDSFDVYSRDTLTIHQDKPALFRSFKWLKPG 360
 DB 322 SPALERSIGLCAVEFEVADCTKINYPDSFDVYSRDTLTIHQDKPALFRSFKWLKPG 381
 QY 361 GKVLSDYCKAGPSPSPFAAYIKRGYDLHDVKEVQOMKDGAGVDVLAERTQPIRV 420
 DB 382 GKVLSDYCKAGPSPSPFAAYIKRGYDLHDVKEVQOMKDGAGVDVLAERTQPIRV 441
 QY 421 LRKELETKEKQVPIIDFSEEDYNDIVGGNNDKLARTKGEORWGLFVAKK 472
 DB 442 LQKELDALKQSKDPIIDFSEEDYNDIVGGNNDKLARTKGEORWGLFVAKK 493

RESULT 2
 PBM2_ARATH STANDARD; PRT: 475 AA.

AC Q94H0; Q9LP63; Q9LP64;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Putative phosphoethanolamine N-methyltransferase 2 (BC 2.1.1.103).
 GN NMT2 OR AT1G48600 OR T1N15.20 OR T1N15.22/T1N15.23.
 OS Arabidopsis thaliana (Muege-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]

SEQUENCE FROM N.A.
 STRAIN=cv. Columbia;
 MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 White O., Alonso J., Arafat H., Araujo R., Bowman C.L., Brooks S.Y.,
 Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
 Millican J., Miranda M., Nguyen M., Mierman W.C., Osborne B.I.,
 Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 Sun H., Tallon L.J., Tabunga G., Toriumi M.J., Town C.D.,
 Utterback T., Van Aken S., Vaysberg M., Vysotskaja V.S., Walker M.,
 Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 *Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.*;

RL Nature 408:816-820(2000).
 RP [2]
 RN SEQUENCE FROM N.A.
 RX STRAIN=cv. Columbia;
 RA MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 Miranda M., Quach H.L., Tripp M., Chang J.M., Lee J.M., Toriumi M.J.,
 Chan M.M., Tang C.C., Onodera C.S., Deng J.H., Brooks S.Y., Carninci P.,
 Arakawa T., Banh J., Banno F., Bowser L., Goldsmith A.D., Gurjal M., Hansen N.F.,
 Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
 Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 Kamiya A., Meyers R., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 Sato M., Tenen R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 Yuan S., Shinozaki K., Davis R.M., Theologis A., Ecker J.R.;
 *Empirical analysis of transcriptional activity in the Arabidopsis
 genome.*;
 RL Science 302:842-846(2003).

CC -!- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,
 phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the
 three methylation steps required to convert phosphoethanolamine to
 phosphocholine (By similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine
 phosphate = S-adenosyl-L-homocysteine + N-methyl ethanolamine
 phosphate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the methyltransferase superfamily.
 CC -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
 gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AC020889; AAF79704.1; ALT SEQ.
 CC EMBL; AC020889; AAF79705.1; ALT SEQ.
 CC EMBL; AF428454; AAL16223.1;
 CC EMBL; AY063866; AAL36222.1;
 CC InterPro; IPR001601; Methyltransf.
 CC InterPro; IPR000051; SAM bind.
 CC InterPro; IPR004033; UbiE/COO5 Metrf.
 CC Pfam; PF01209; Ubie methyltran; 1.
 CC Transferase; Methyltransferase; Repeat.
 CC DOMAIN 34 143 SAM-BINDING 1.
 CC FT DOMAIN 263 370 SAM-BINDING 2.
 CC SEQUENCE 475 AA; 54018 MW; 04070544DE80DA8 CRC64;

Query Match 82.6%; Score 2077; DB 1; Length 475;
 Best Local Similarity 80.1%; Pred. No. 3.2e-149;
 Matches 379; Conservative 47; Mismatches 47; Indels 0; Gaps 0;

QY 1 HTVDLTTTAAWMLDSQASDLKBERPHTLSMLPPELGKCLLELGAGIGRTGELAAGQOV 60
 DB 3 HSDVDTVEAWMLDSQASDLKBERPHTLSMLPPELGKCLLELGAGIGRTGELAAGQOV 62
 QY 61 IALDPIESAIKNEVINGHYKVKFPCADVTSTLSPFPHSDLVFNSWLLMYLSDEVE 120
 DB 63 IALDPIESAIKNEVINGHYKVKFPCADVTSTLSPFPHSDLVFNSWLLMYLSDEVE 122
 QY 121 NLVERMLKWLKPGGYIFFRESCHQSDHKKSNPTHYREPRFTYKAPKCECHLQDGSNS 180
 DB 123 LRVEMMLKWLKPGGYIFFRESCHQSDHKKSNPTHYREPRFTYKAPKCECHLQDGSNS 182
 QY 181 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSQYKCNLSLYRYVP 240
 DB 183 YELSLSCCKTGAVYKVKNNQNOISLWKQVDSKDKGFORPLDTSQYKCNLSLYRYVP 242

241 GPGVSTGGYETTKFVSMMLDKPGQKVLVDGCGGDFMAETFDVVGFDLSVNNI 300
 243 GEGVSTGGYETTKFVSMMLDKPGQKVLVDGCGGDFMAETFDVVGFDLSVNNI 302
 301 SPALERSIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKPAFRSPYKWLKRG 360
 303 SPALERSIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKPAFRSPYKWLKRG 362
 361 GKVLISDYCKGAGPSPFAAYIKRGYDLHDVKEYGMLKADGPDVLAORTQPIRV 420
 363 GKVLITDYCRSAETSPFAAYIKRGYDLHDVKEYGMLKADGPDVLAORTQPIRV 422
 421 LRKELETVEKEDVTSFSDSEYNDIVGWNDRKRTAKGEORGLFVAKK 473
 423 LRKELEKVEKEEFISFSDSEYNDIVGWNDRKRTAKGEORGLFVAKK 475

RESULT 3

PEM1 ARATH STANDARD; PRT; 490 AA.
 AC Q9C6B9; Q9C9V1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative phosphoethanolamine N-methyltransferase 3 (EC 2.1.1.103).
 GN NMT3 OR AT1G73600 OR P6D5.1 OR P25P22.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsi.
 NCBI_TaxID=3702;
 RS [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Fedorapeli N.A., Kaul S.,
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler S., Chan A., Chao Q., Chen H., Cheuk R.P., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin B.,
 RA Kim C.J., Koo H.L., Krenetsaka I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.A., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Millican J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vyeotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 FT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:816-820(2000).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Schneider M.;
 CC Unpublished observations (MAY-2002).
 CC -1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,
 CC phosphomonomethyl ethanolamine and phosphodimethyl ethanolamine, the
 CC three methylation steps required to convert phosphoethanolamine to
 CC phosphocholine (By similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine
 CC phosphate = S-adenosyl-L-homocysteine + N-methyl ethanolamine
 CC phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the methyltransferase superfamily.
 CC -1- CAUTION: Ref.1 (AGS1806) sequence differs from that shown due to
 CC erroneous gene model prediction.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

DR EMBL; AC079676; AAG51806.1; ALT SEQ.
 DR EMBL; AC012679; AAG52075.1; -
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 KM Transferase; Methyltransferase; Repeat.
 FT DOMAIN 49 158 SAM-BINDING 2.
 FT DOMAIN 278 385 SAM-BINDING 2.
 SQ SEQUENCE 490 AA; 56368 MW; 77DFAFBC89C41CB CRC64;

Query Match 81.2%; Score 2041; DB 1; Length 490;
 Best Local Similarity 78.4%; Pred. No. 1.7e-146;
 Matches 370; Conservative 53; Mismatches 49; Indels 0; Gaps 0;

OY 1 HTVDLTIANMMLDSQASDLDEXEPRLSMLPPEKCKLLELGAIGRTGLAKAGOV 60
 DB 18 HVGSLSEAMMLDSKASDLDEXEPRLAFLPPIEGTTLFEGAGIGRTTFLAQKAGOV 77
 OY 61 IALDPTESAIKKNVINGHYKVKFEMCADVTSTLSPPHSLDVIFSNWLLMYLSDEVE 120
 DB 78 IAVDFIESVIKKNENINGHYKVKFELCDVTSNNMFPNESMDLIFSNWLLMYLSDEVE 137
 OY 121 NIVEMMLKWLKGGYIPFSCFHSQSDHQRKSNPTHYRPRPTTKAFKECHLDQSGNS 180
 DB 138 DLAKMQLQWTVGGYIPFSCFHSQSDHQRKSNPTHYRPRPTTKAFKECHMDNEDGNS 197
 OY 181 YELSLSCCKCTGAYVRNKNQNLISLWQKVDSDKQFORFLDTSTQYKNSILYERVF 240
 DB 198 YELSLVSCCKCIGAYVRNKNQNLISLWQKVDSDKQFORFLDTSTQYKNSILYERVF 257
 OY 241 GCGYSTGCTGCTTGFVSMMLDKPGQKVLVDGCGGDFMAETFDVVGFDLSVNNI 300
 DB 258 GCGYSTGCTGCTTGFVSMMLDKPGQKVLVDGCGGDFMAETFDVVGFDLSVNNI 317
 OY 301 SPALERSIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKPAFRSPYKWLKRG 360
 DB 318 SPALERSIGLCAVSEFVADCTKINYPNSDVTYSRDTLHODKPAFRSPYKWLKRG 377
 OY 361 GKVLISDYCKGAGPSPFAAYIKRGYDLHDVKEYGMLKADGPDVLAORTQPIRV 420
 DB 378 GKVLITDYCRSAETSPFAAYIKRGYDLHDVKEYGMLKADGPDVLAORTQPIRV 422
 OY 421 LRKELETVEKEDVTSFSDSEYNDIVGWNDRKRTAKGEORGLFVAKK 472
 DB 438 LRKELEKVEKEEFISFSDSEYNDIVGWNDRKRTAKGEORGLFVAKK 475

RESULT 4

PEM1 ARATH STANDARD; PRT; 491 AA.
 ID PEM1 ARATH STANDARD; PRT; 491 AA.
 AC Q9FR44; Q9LVH3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Phosphoethanolamine N-methyltransferase 1 (EC 2.1.1.103) (PEAMT 1)
 DE (AtNMT1).
 GN NMT1 OR AT3G17990/AT3G18000 OR MEBS.19 OR MEBS.21/MEBS.22.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta; TIGR=Seedling;
 RX MEDLINE=20567827; PubMed=11115895;
 RA Bolognese C.P., McGraw P.;
 FT "The isolation and characterization in yeast of a gene for Arabidopsis
 S-adenosylmethionine:phospho-ethanolamine N-methyltransferase.";

RL	Plant Physiol. 124:1800-1813 (2000).	61	I	ALDIESAIKXNEVINGHYKXKFCADVTSPILSPHSLDVI	ESWLLMYLSDEVE	120
RN	SEQUENCE FROM N.A.	62	:			121
RP	STRAIN=cv. Columbia;	63	:			122
RC	MEDLINE=20277480; PubMed=10819329;	64	:			123
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;	65	:			124
RA	"Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence	66	:			125
RT	features of the regions of 4,504,864 bp covered by sixty P1 and TAC	67	:			126
RT	clones"; 7:131-135 (2000).	68	:			127
RT	DNA Res. 7:131-135 (2000).	69	:			128
RN	SEQUENCE FROM N.A.	70	:			129
RP	STRAIN=cv. Columbia;	71	:			130
RC	MEDLINE=22954850; PubMed=14593172;	72	:			131
RA	Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Cheuk R.P.,	73	:			132
RA	Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Yu G.,	74	:			133
RA	Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,	75	:			134
RA	Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,	76	:			135
RA	Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,	77	:			136
RA	Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,	78	:			137
RA	Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,	79	:			138
RA	Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,	80	:			139
RA	Khan S., Koesema E., Iehida J., Jiang P.X., Jones T., Kawai J.,	81	:			140
RA	Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,	82	:			141
RA	Satou M., Tamse R., Vayenberg N., Wallender E.K., Wong C., Yamamura Y.,	83	:			142
RA	Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;	84	:			143
RT	"Empirical analysis of transcriptional activity in the Arabidopsis	85	:			144
RT	genome";	86	:			145
RL	Science 302:842-846 (2003).	87	:			146
CC	-1- FUNCTION: Catalyzes N-methylation of phosphoethanolamine,	88	:			147
CC	phosphonomethyl ethanolamine and phosphodimethyl ethanolamine, the	89	:			148
CC	three methylation steps required to convert phosphoethanolamine to	90	:			149
CC	phosphocholine.	91	:			150
CC	-1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + ethanolamine	92	:			151
CC	phosphate = S-adenosyl-L-homocysteine + N-methylethanolamine	93	:			152
CC	phosphate.	94	:			153
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).	95	:			154
CC	-1- SIMILARITY: Belongs to the methyltransferase superfamily.	96	:			155
CC	-1- CAUTION: This protein is a fusion of the two annotated genes	97	:			156
CC	At3g17990 and At3g18000.	98	:			157
CC	-1- CAUTION: Ref.2 sequence differs from that shown due to erroneous	99	:			158
CC	gene model prediction.	100	:			159
CC	-----	101	:			160
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-	102	:			161
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstation	103	:			162
CC	at the European Bioinformatics Institute. There are no restrictions on its	104	:			163
CC	use by non-profit institutions as long as its content is in no way	105	:			164
CC	modified and this statement is not removed. Usage by and for commercial	106	:			165
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	107	:			166
CC	or send an email to license@sib-sib.ch).	108	:			167
CC	-----	109	:			168
DR	EMBL; AF197940; AAC41121.1;	110	:			169
DR	EMBL; AB019230; BAB02720.1; ALT_SEQ.	111	:			170
DR	EMBL; AF367299; AAK32886.1;	112	:			171
DR	EMBL; AY058175; AAL25589.1;	113	:			172
DR	EMBL; AY091683; AAM10282.1;	114	:			173
DR	EMBL; AY093093; AAM13092.1;	115	:			174
DR	InterPro; IPR001601; Methyltransf.	116	:			175
DR	InterPro; IPR00051; SAM bind.	117	:			176
KW	Methyltransferase; Transferase; Repeat.	118	:			177
PT	DOMAIN 50 159 SAM-BINDING 1.	119	:			178
PT	DOMAIN 279 386 SAM-BINDING 2.	120	:			179
PT	CONFLICT 333 333 E -> G (IN REF. 3; AAM13092).	121	:			180
SQ	SEQUENCE 451 AA; 56102 MW; DA4404EBED3FABD5 CRC64;	122	:			181
Query Match	80.0%; Score 2011; DB 1; Length 491;	123	:			182
Best Local Similarity	77.8%; Pred. No. 3.2e-144;	124	:			183
Matches 367; Conservative	51; Mismatches 54; Indels	125	:			184
0; Gaps	0;	126	:			185
1	HTVLTIRAWLDQASDLDRPETHLNLPLPEKCLLELGACIGFTGELAKAGQV. 6Q	127	:			186
19	HSADLTVENWLDLRASDLDRPETHLNLPLPEKCLLELGACIGFTGELAKAGEL. 78	128	:			187
CC	-----	129	:			188
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-	130	:			189

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE015852; AAN57171.1; -
 DR HAMAP; MP_01813; -; 1.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 DR InterPro; IPR004034; UniProt/MetTransf.
 DR InterPro; IPR004033; UbiE/COG5_MetTrf.
 DR Pfam; PF01209; UbiE_methyltransf; 1.
 DR PROSITE; PS01183; UBI1_1; 1.
 DR PROSITE; PS01184; UBI2_1; 1.
 KW Menquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
 KW Methyltransferase; Complete proteome.
 SQ SEQUENCE 251 AA; 28065 MW; EPOAFC3CF06B2141 CRC64;

Query Match 7.0%; Score 176.5; DB 1; Length 251;
 Best Local Similarity 28.3%; Pred. No. 3.5e-06;
 Matches 63; Conservative 43; Mismatches 76; Indels 41; Gaps 12;

OY 245 VSTGQYETKTF-VMSLDLKPQKQVDPVGGIGG-----GDFYMAETPDVWVGFD 294
 DB 43 MSFGIHRWRYTTEVSGARPGMKVLDLAGTGLTAKFSLVGD-----KGEVVLAD 95
 OY 295 LSVNMISFALP-RSIGLKCAVEPADCTKINYPDNPSPDVIYSRDTLHIDQKALPRS 352
 DB 96 INDSMLKVGRTKLRDGIIVSNVSVQANAEALFPDNDHFDIITIAFGLNVTDKDALRS 155
 OY 353 FYKMLKPGKGLVLDYCKKAGPSPEPAVYKQGVLDHVK---EYGM-L-KDAGFVDV 408
 DB 156 MVRVLPKPGKGLVLFESK-----PQHE-----VMKVVDLYSFVLPQMGQLITKQADSYEY 207
 OY 409 LAEDTEGFIVLRKLETVKEKQVFI-SDFSEEDYNDIVGG 450
 DB 208 LAES-----IR-MHPDDDTL---KQMVVDAGFEQVDYTNNTDG 241

RESULT 6
 GTOM ARATH STANDARD; PRT; 348 AA.
 AC Q2SK1; Q9XIP9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Tocopherol O-methyltransferase, chloroplast precursor (EC 2.1.1.95)
 DE [Gamma-tocopherol methyltransferase].
 GN G-TMT OR AT1G64970 OR F13011.27.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopses.
 OX NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Shintani D.; DellaPenna D.;
 RX MEDLINE=99069631; PubMed=9851934;
 RT "Elevating the vitamin E content of plants through metabolic
 engineering";
 RL Science 283:2098-2100(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA White O.; Alonso J.; Altariff H.; Araujo R.; Bowman C.L.; Brooks S.Y.;
 RA Buehler E.; Chan A.; Chao Q.; Chen H.; Cheuk R.P.; Chin C.W.;
 RA Chung M.K.; Conn L.; Conway A.B.; Conway A.R.; Creasy T.H.; Dewar K.;
 RA Dunn P.; Egn P.; Feldblum T.V.; Feng J.-D.; Pong B.; Fujii C.Y.;

RA Gill J.E.; Goldsmith A.D.; Haas B.; Hansen N.F.; Hughes B.; Huizar L.;
 RA Hunter J.L.; Jenkins J.; Johnson-Hopson C.; Khan S.; Khaykin E.;
 RA Kim C.J.; Koo H.B.; Kremenetskaia I.; Kurtz D.B.; Kvan A.; Lam B.;
 RA Langin-Hooper S.; Lee A.; Lee J.M.; Lenz C.A.; Li J.H.; Li Y.-P.;
 RA Lin X.; Liu S.X.; Liu Z.A.; Lueros J.S.; Maiti R.; Marziani A.;
 RA Militscher J.; Miranda M.; Nguyen M.; Nierman W.C.; Osborne B.I.;
 RA Pai G.; Peterson J.; Pham P.K.; Rizzo M.; Rooney T.; Rowley D.;
 RA Sakano H.; Salzberg S.L.; Schwartz J.R.; Shinn P.; Southwick A.M.;
 RA Sun H.; Tallon L.J.; Tambunga G.; Toriumi M.J.; Town C.D.;
 RA Utterback T.; Van Aken S.; Vaysberg M.; Vysotskaia V.S.; Walker M.;
 RA Wu D.; Yu G.; Fraser C.M.; Venter J.C.; Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K.; Lim J.; Dale J.M.; Chen H.; Shinn P.; Palm C.J.;
 RA Southwick A.M.; Wu H.C.; Kim C.J.; Nguyen M.; Pham P.K.; Cheuk R.F.;
 RA Karlin-Newmann G.; Liu S.X.; Lam B.; Sakano H.; Wu T.; Yu G.;
 RA Miranda M.; Quach H.L.; Tripp M.; Chang C.H.; Lee J.M.; Toriumi M.J.;
 RA Chan M.T.; Tang C.C.; Banno F.; Bowser L.; Brooks S.Y.; Carninci P.;
 RA Arakawa T.; Banh J.; Banno F.; Deng J.M.; Akiyama K.; Ansari Y.;
 RA Chao Q.; Choy N.; Enju A.; Goldsmith A.D.; Gurjal M.; Hansen N.F.;
 RA Hayashizaki Y.; Johnson-Hopson C.; Hsuan V.W.; Iida K.; Karnes M.;
 RA Khan S.; Koesema B.; Ishida J.; Jiang P.X.; Jones T.; Kawai J.;
 RA Kamiya A.; Meyers C.; Nakajima M.; Narusaka M.; Seki M.; Sakurai T.;
 RA Satou M.; Tanase R.; Vaysberg M.; Wallender E.K.; Wong C.; Yamamura Y.;
 RA Yuan S.; Shinozaki K.; Davis R.W.; Theologis A.; Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 genome";
 RL Science 302:842-846(2003).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Brover V.; Troukhan M.; Alexandrov N.; Lu Y.-P.; Flavell R.;
 RA Feldmann K.A.;
 RT "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Methylates gamma- and delta-tocopherol to yield beta-
 and alpha-tocopherol respectively.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + gamma-tocopherol =
 S-adenosyl-L-homocysteine + alpha-tocopherol.
 CC -!- PATHWAY: Vitamin E biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF104220; AAD02882.1; -
 DR EMBL; AC006193; AAD38271.2; -
 DR EMBL; AY049258; AAK83600.1; -
 DR EMBL; AY090280; AAL90941.1; -
 DR EMBL; AY087138; AAM64696.1; -
 DR PIR; C96673; C96673.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 KW Methyltransferase; Transferase; Chloroplast; Transit peptide.
 FT TRANSIT 1 50 CHLOROPLAST (POTENTIAL).
 FT CHAIN 51 348 TOCOPHEROL O-METHYLTRANSFERASE.
 FT DOMAIN 9 40 SER-RICH.
 FT DOMAIN 123 232 SAM-BINDING.
 FT CONFLICT 175 175 S -> A (IN REF. 2 AND 3).
 FT CONFLICT 188 188 Q -> K (IN REF. 4).
 SQ SEQUENCE 348 AA; 38091 MW; 47DB18A72E075F CRC64;

Query Match 6.9%; Score 173.5; DB 1; Length 348;
 Best Local Similarity 27.2%; Pred. No. 9.1e-06;

Matches 52; Conservative 30; Mismatches 64; Indels 45; Gaps 7;

QY 202 NOISLWQKV-----DSK-----DDKGFQ-----RFLDTQYKCNSTLRVERVFGP 242

Db 69 NETSGLWEEIWDHMHGFDYDPSSVQLSDSGHKEAQIMIES-----LRPAGV--- 118

QY 243 GYVSTGGYETTFEVSMLDKPKQKVLVDCGIGGDDFYMAETFDVVVGGFSLVNMISP 302

Db 119 -----TDEE-----BEKTKKVVWVCGCGSSRYLAKPGACIGITLSPVQAKR 164

QY 303 A--LERSIGLKCAREFEVADCTKINYPDNSFDVIYSRDTILHIQKPALPRSPYKMLKPG 360

Db 165 ANDLAAQSLSHKASFOVADALQPPEDGKFDLVMSBSGEHMPKAFVKVELVRAVAP 224

QY 361 GNVLSDYCKK 371

Db 225 GRILLVWCHR 235

RESULT 7

UBIE STAAAM STANDARD; PRT; 241 AA.

AC Q99U19;

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Menquinone biosynthesis methyltransferase ubiB (EC 2.1.1.-).

GN UBI8 OR SAV1471 OR SNA1303 OR MW1360.

OS Staphylococcus aureus (strain Mu50 / ATCC 700699).

OS Staphylococcus aureus (strain N315), and

OS Staphylococcus aureus (strain MW2).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878, 158879, 196620;

RN SEQUENCE FROM N.A.

RP STRAIN=MU50 / ATCC 700699, and N315;

RX MEDLINE=21311952; PubMed=11418146;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

RT aureus".

RL Lancet 357:1225-1240(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MW2;

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,

RA Yamamoto K., Hiramatsu K.;

RT "Genome and virulence determinants of high virulence community-

RT acquired MRSA".

RL Lancet 359:1819-1827(2002).

CC -1- FUNCTION: Methyltransferase required for the conversion of

CC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) (By similarity).

CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =

CC S-adenosyl-L-homocysteine + menaquinol.

CC -1- PATHWAY: Menaquinone biosynthesis; last step.

CC -1- SIMILARITY: Belongs to the ubiE family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. (See <http://www.ebi.ac.uk/announcements/>

CC or send an email to license@ebi.ac.uk.)

CC -----

EMBL; AP003362; BAB57633.1; -.

DR EMBL; AP003362; BAB57633.1; -.

DR EMBL; AP004827; BAB95225.1; -.

DR PIR; G89925; G89925.

DR HAMAP; ME 101813; -; 1.

DR InterPro; IPR001601; Methyltransf.

DR InterPro; IPR000051; SAM_bind.

DR InterPro; IPR004034; Ubi7men_Mettransf.

DR InterPro; IPR004033; UbiE/COG5_Metrf.

DR Pfam; PF01209; UbiE_methyltran_1

DR PROSITE; PS01184; UBI1_1; FALSE_NEG.

DR PROSITE; PS01184; UBI1_2; 1.

KW Menaquinone biosynthesis; Transferase; Methyltransferase;

KW Complete proteome.

SQ SEQUENCE 241 AA; 27423 MW; 90FBF4020ABCD54 CRC64;

Query Match 6.7%; Score 167.5; DB 1; Length 241;

Best Local Similarity 23.0%; Pred. No. 1.6e-05;

Matches 56; Conservative 46; Mismatches 79; Indels 63; Gaps 8;

QY 197 NKKVQNIQISMLWQVDSKDKGFQFLDTQYKCNSTLRVERVFGP 254

Db 4 NKAKEQVHVFQNIKKYD-----ALNNIISFEQHKVW-----RK 39

QY 255 EFSKMLDLKPKQKVLVDCGIGGDDFYMAETFDV-----EYVGFSLVNMISPALERSIGL 310

Db 40 RVKMDGMYRKGTKALDVCCGT--GDWTIALSKAVGTEVTGIDFSENMLVGVGKEKTASM 97

QY 311 KCAVEFEVADCTKINYPDNSFDVIYSRDTILHIQKPALPRSPYKMLKPGKVLISDYCK 370

Db 98 E-NVKLVHGDAMLPEDNSFDVYIGFGLRNVPDYLAKEMNRVLKPGQWV---CL 152

QY 371 KAGPP-----SPEPAAYIKORGYDLHDVKEVGOMLKDAG 404

Db 153 ETSQPTLPVFKQYALYFKVFMIFGKLPAKSKRYENLQOSTFAPFGKSELKRNFBAG 212

QY 405 FVDV 408

Db 213 FINV 216

RESULT 8

UBIE VIBCH STANDARD; PRT; 260 AA.

AC Q9KV06;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiB

DE (EC 2.1.1.-).

GN UBI8 OR VC0083.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrrio.

OX NCBI_TaxID=6661;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=El Tor N16961 / Serotype O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,

RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,

RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Dragoi I., Sellers P.,

RA Smolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,

RA McDonald L., Utterback T., Fleischmann R.D., Niemann W.C., White O.,

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

RT cholerae".

RL Nature 406:477-483(2000).

CC -1- FUNCTION: Methyltransferase required for the conversion of

CC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the

CC conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to

CC 2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DDMQH2) (By

CC similarity).

CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyphenyl-6-
 CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
 CC polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
 CC s-adenosyl-L-homocysteine + menaquinol.
 CC -1- PATHWAY: Menaquinone biosynthesis; last step.
 CC -1- PATHWAY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ubiE family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE004099; AAF93261.1; -
 CC PIR; D82366; D82366.
 CC TIGR; VC0083; -
 CC HAMAP; MF_01813; -; 1.
 CC InterPro; IPR001601; Methyltransf.
 CC InterPro; IPR000051; SAM bind.
 CC InterPro; IPR004034; Ubi7/en Mentransf.
 CC InterPro; IPR004033; UbiE/COGS_Metrif.
 CC Pfam; PF01209; UbiE_methyltrans; 1.
 CC PROSITE; PS01183; UBI1; 1.
 CC PROSITE; PS01184; UBI2; 1.
 CC K01183; UBI1; 1.
 CC K01184; UBI2; 1.
 CC K01185; UBI3; 1.
 CC K01186; UBI4; 1.
 CC K01187; UBI5; 1.
 CC K01188; UBI6; 1.
 CC K01189; UBI7; 1.
 CC K01190; UBI8; 1.
 CC K01191; UBI9; 1.
 CC K01192; UBI10; 1.
 CC K01193; UBI11; 1.
 CC K01194; UBI12; 1.
 CC K01195; UBI13; 1.
 CC K01196; UBI14; 1.
 CC K01197; UBI15; 1.
 CC K01198; UBI16; 1.
 CC K01199; UBI17; 1.
 CC K01200; UBI18; 1.
 CC K01201; UBI19; 1.
 CC K01202; UBI20; 1.
 CC K01203; UBI21; 1.
 CC K01204; UBI22; 1.
 CC K01205; UBI23; 1.
 CC K01206; UBI24; 1.
 CC K01207; UBI25; 1.
 CC K01208; UBI26; 1.
 CC K01209; UBI27; 1.
 CC K01210; UBI28; 1.
 CC K01211; UBI29; 1.
 CC K01212; UBI30; 1.
 CC K01213; UBI31; 1.
 CC K01214; UBI32; 1.
 CC K01215; UBI33; 1.
 CC K01216; UBI34; 1.
 CC K01217; UBI35; 1.
 CC K01218; UBI36; 1.
 CC K01219; UBI37; 1.
 CC K01220; UBI38; 1.
 CC K01221; UBI39; 1.
 CC K01222; UBI40; 1.
 CC K01223; UBI41; 1.
 CC K01224; UBI42; 1.
 CC K01225; UBI43; 1.
 CC K01226; UBI44; 1.
 CC K01227; UBI45; 1.
 CC K01228; UBI46; 1.
 CC K01229; UBI47; 1.
 CC K01230; UBI48; 1.
 CC K01231; UBI49; 1.
 CC K01232; UBI50; 1.
 CC K01233; UBI51; 1.
 CC K01234; UBI52; 1.
 CC K01235; UBI53; 1.
 CC K01236; UBI54; 1.
 CC K01237; UBI55; 1.
 CC K01238; UBI56; 1.
 CC K01239; UBI57; 1.
 CC K01240; UBI58; 1.
 CC K01241; UBI59; 1.
 CC K01242; UBI60; 1.
 CC K01243; UBI61; 1.
 CC K01244; UBI62; 1.
 CC K01245; UBI63; 1.
 CC K01246; UBI64; 1.
 CC K01247; UBI65; 1.
 CC K01248; UBI66; 1.
 CC K01249; UBI67; 1.
 CC K01250; UBI68; 1.
 CC K01251; UBI69; 1.
 CC K01252; UBI70; 1.
 CC K01253; UBI71; 1.
 CC K01254; UBI72; 1.
 CC K01255; UBI73; 1.
 CC K01256; UBI74; 1.
 CC K01257; UBI75; 1.
 CC K01258; UBI76; 1.
 CC K01259; UBI77; 1.
 CC K01260; UBI78; 1.
 CC K01261; UBI79; 1.
 CC K01262; UBI80; 1.
 CC K01263; UBI81; 1.
 CC K01264; UBI82; 1.
 CC K01265; UBI83; 1.
 CC K01266; UBI84; 1.
 CC K01267; UBI85; 1.
 CC K01268; UBI86; 1.
 CC K01269; UBI87; 1.
 CC K01270; UBI88; 1.
 CC K01271; UBI89; 1.
 CC K01272; UBI90; 1.
 CC K01273; UBI91; 1.
 CC K01274; UBI92; 1.
 CC K01275; UBI93; 1.
 CC K01276; UBI94; 1.
 CC K01277; UBI95; 1.
 CC K01278; UBI96; 1.
 CC K01279; UBI97; 1.
 CC K01280; UBI98; 1.
 CC K01281; UBI99; 1.
 CC K01282; UBI100; 1.
 CC K01283; UBI101; 1.
 CC K01284; UBI102; 1.
 CC K01285; UBI103; 1.
 CC K01286; UBI104; 1.
 CC K01287; UBI105; 1.
 CC K01288; UBI106; 1.
 CC K01289; UBI107; 1.
 CC K01290; UBI108; 1.
 CC K01291; UBI109; 1.
 CC K01292; UBI110; 1.
 CC K01293; UBI111; 1.
 CC K01294; UBI112; 1.
 CC K01295; UBI113; 1.
 CC K01296; UBI114; 1.
 CC K01297; UBI115; 1.
 CC K01298; UBI116; 1.
 CC K01299; UBI117; 1.
 CC K01300; UBI118; 1.
 CC K01301; UBI119; 1.
 CC K01302; UBI120; 1.
 CC K01303; UBI121; 1.
 CC K01304; UBI122; 1.
 CC K01305; UBI123; 1.
 CC K01306; UBI124; 1.
 CC K01307; UBI125; 1.
 CC K01308; UBI126; 1.
 CC K01309; UBI127; 1.
 CC K01310; UBI128; 1.
 CC K01311; UBI129; 1.
 CC K01312; UBI130; 1.
 CC K01313; UBI131; 1.
 CC K01314; UBI132; 1.
 CC K01315; UBI133; 1.
 CC K01316; UBI134; 1.
 CC K01317; UBI135; 1.
 CC K01318; UBI136; 1.
 CC K01319; UBI137; 1.
 CC K01320; UBI138; 1.
 CC K01321; UBI139; 1.
 CC K01322; UBI140; 1.
 CC K01323; UBI141; 1.
 CC K01324; UBI142; 1.
 CC K01325; UBI143; 1.
 CC K01326; UBI144; 1.
 CC K01327; UBI145; 1.
 CC K01328; UBI146; 1.
 CC K01329; UBI147; 1.
 CC K01330; UBI148; 1.
 CC K01331; UBI149; 1.
 CC K01332; UBI150; 1.
 CC K01333; UBI151; 1.
 CC K01334; UBI152; 1.
 CC K01335; UBI153; 1.
 CC K01336; UBI154; 1.
 CC K01337; UBI155; 1.
 CC K01338; UBI156; 1.
 CC K01339; UBI157; 1.
 CC K01340; UBI158; 1.
 CC K01341; UBI159; 1.
 CC K01342; UBI160; 1.
 CC K01343; UBI161; 1.
 CC K01344; UBI162; 1.
 CC K01345; UBI163; 1.
 CC K01346; UBI164; 1.
 CC K01347; UBI165; 1.
 CC K01348; UBI166; 1.
 CC K01349; UBI167; 1.
 CC K01350; UBI168; 1.
 CC K01351; UBI169; 1.
 CC K01352; UBI170; 1.
 CC K01353; UBI171; 1.
 CC K01354; UBI172; 1.
 CC K01355; UBI173; 1.
 CC K01356; UBI174; 1.
 CC K01357; UBI175; 1.
 CC K01358; UBI176; 1.
 CC K01359; UBI177; 1.
 CC K01360; UBI178; 1.
 CC K01361; UBI179; 1.
 CC K01362; UBI180; 1.
 CC K01363; UBI181; 1.
 CC K01364; UBI182; 1.
 CC K01365; UBI183; 1.
 CC K01366; UBI184; 1.
 CC K01367; UBI185; 1.
 CC K01368; UBI186; 1.
 CC K01369; UBI187; 1.
 CC K01370; UBI188; 1.
 CC K01371; UBI189; 1.
 CC K01372; UBI190; 1.
 CC K01373; UBI191; 1.
 CC K01374; UBI192; 1.
 CC K01375; UBI193; 1.
 CC K01376; UBI194; 1.
 CC K01377; UBI195; 1.
 CC K01378; UBI196; 1.
 CC K01379; UBI197; 1.
 CC K01380; UBI198; 1.
 CC K01381; UBI199; 1.
 CC K01382; UBI200; 1.
 CC K01383; UBI201; 1.
 CC K01384; UBI202; 1.
 CC K01385; UBI203; 1.
 CC K01386; UBI204; 1.
 CC K01387; UBI205; 1.
 CC K01388; UBI206; 1.
 CC K01389; UBI207; 1.
 CC K01390; UBI208; 1.
 CC K01391; UBI209; 1.
 CC K01392; UBI210; 1.
 CC K01393; UBI211; 1.
 CC K01394; UBI212; 1.
 CC K01395; UBI213; 1.
 CC K01396; UBI214; 1.
 CC K01397; UBI215; 1.
 CC K01398; UBI216; 1.
 CC K01399; UBI217; 1.
 CC K01400; UBI218; 1.
 CC K01401; UBI219; 1.
 CC K01402; UBI220; 1.
 CC K01403; UBI221; 1.
 CC K01404; UBI222; 1.
 CC K01405; UBI223; 1.
 CC K01406; UBI224; 1.
 CC K01407; UBI225; 1.
 CC K01408; UBI226; 1.
 CC K01409; UBI227; 1.
 CC K01410; UBI228; 1.
 CC K01411; UBI229; 1.
 CC K01412; UBI230; 1.
 CC K01413; UBI231; 1.
 CC K01414; UBI232; 1.
 CC K01415; UBI233; 1.
 CC K01416; UBI234; 1.
 CC K01417; UBI235; 1.
 CC K01418; UBI236; 1.
 CC K01419; UBI237; 1.
 CC K01420; UBI238; 1.
 CC K01421; UBI239; 1.
 CC K01422; UBI240; 1.
 CC K01423; UBI241; 1.
 CC K01424; UBI242; 1.
 CC K01425; UBI243; 1.
 CC K01426; UBI244; 1.
 CC K01427; UBI245; 1.
 CC K01428; UBI246; 1.
 CC K01429; UBI247; 1.
 CC K01430; UBI248; 1.
 CC K01431; UBI249; 1.
 CC K01432; UBI250; 1.
 CC K01433; UBI251; 1.
 CC K01434; UBI252; 1.
 CC K01435; UBI253; 1.
 CC K01436; UBI254; 1.
 CC K01437; UBI255; 1.
 CC K01438; UBI256; 1.
 CC K01439; UBI257; 1.
 CC K01440; UBI258; 1.
 CC K01441; UBI259; 1.
 CC K01442; UBI260; 1.
 CC K01443; UBI261; 1.
 CC K01444; UBI262; 1.
 CC K01445; UBI263; 1.
 CC K01446; UBI264; 1.
 CC K01447; UBI265; 1.
 CC K01448; UBI266; 1.
 CC K01449; UBI267; 1.
 CC K01450; UBI268; 1.
 CC K01451; UBI269; 1.
 CC K01452; UBI270; 1.
 CC K01453; UBI271; 1.
 CC K01454; UBI272; 1.
 CC K01455; UBI273; 1.
 CC K01456; UBI274; 1.
 CC K01457; UBI275; 1.
 CC K01458; UBI276; 1.
 CC K01459; UBI277; 1.
 CC K01460; UBI278; 1.
 CC K01461; UBI279; 1.
 CC K01462; UBI280; 1.
 CC K01463; UBI281; 1.
 CC K01464; UBI282; 1.
 CC K01465; UBI283; 1.
 CC K01466; UBI284; 1.
 CC K01467; UBI285; 1.
 CC K01468; UBI286; 1.
 CC K01469; UBI287; 1.
 CC K01470; UBI288; 1.
 CC K01471; UBI289; 1.
 CC K01472; UBI290; 1.
 CC K01473; UBI291; 1.
 CC K01474; UBI292; 1.
 CC K01475; UBI293; 1.
 CC K01476; UBI294; 1.
 CC K01477; UBI295; 1.
 CC K01478; UBI296; 1.
 CC K01479; UBI297; 1.
 CC K01480; UBI298; 1.
 CC K01481; UBI299; 1.
 CC K01482; UBI300; 1.
 CC K01483; UBI301; 1.
 CC K01484; UBI302; 1.
 CC K01485; UBI303; 1.
 CC K01486; UBI304; 1.
 CC K01487; UBI305; 1.
 CC K01488; UBI306; 1.
 CC K01489; UBI307; 1.
 CC K01490; UBI308; 1.
 CC K01491; UBI309; 1.
 CC K01492; UBI310; 1.
 CC K01493; UBI311; 1.
 CC K01494; UBI312; 1.
 CC K01495; UBI313; 1.
 CC K01496; UBI314; 1.
 CC K01497; UBI315; 1.
 CC K01498; UBI316; 1.
 CC K01499; UBI317; 1.
 CC K01500; UBI318; 1.
 CC K01501; UBI319; 1.
 CC K01502; UBI320; 1.
 CC K01503; UBI321; 1.
 CC K01504; UBI322; 1.
 CC K01505; UBI323; 1.
 CC K01506; UBI324; 1.
 CC K01507; UBI325; 1.
 CC K01508; UBI326; 1.
 CC K01509; UBI327; 1.
 CC K01510; UBI328; 1.
 CC K01511; UBI329; 1.
 CC K01512; UBI330; 1.
 CC K01513; UBI331; 1.
 CC K01514; UBI332; 1.
 CC K01515; UBI333; 1.
 CC K01516; UBI334; 1.
 CC K01517; UBI335; 1.
 CC K01518; UBI336; 1.
 CC K01519; UBI337; 1.
 CC K01520; UBI338; 1.
 CC K01521; UBI339; 1.
 CC K01522; UBI340; 1.
 CC K01523; UBI341; 1.
 CC K01524; UBI342; 1.
 CC K01525; UBI343; 1.
 CC K01526; UBI344; 1.
 CC K01527; UBI345; 1.
 CC K01528; UBI346; 1.
 CC K01529; UBI347; 1.
 CC K01530; UBI348; 1.
 CC K01531; UBI349; 1.
 CC K01532; UBI350; 1.
 CC K01533; UBI351; 1.
 CC K01534; UBI352; 1.
 CC K01535; UBI353; 1.
 CC K01536; UBI354; 1.
 CC K01537; UBI355; 1.
 CC K01538; UBI356; 1.
 CC K01539; UBI357; 1.
 CC K01540; UBI358; 1.
 CC K01541; UBI359; 1.
 CC K01542; UBI360; 1.
 CC K01543; UBI361; 1.
 CC K01544; UBI362; 1.
 CC K01545; UBI363; 1.
 CC K01546; UBI364; 1.
 CC K01547; UBI365; 1.
 CC K01548; UBI366; 1.
 CC K01549; UBI367; 1.
 CC K01550; UBI368; 1.
 CC K01551; UBI369; 1.
 CC K01552; UBI370; 1.
 CC K01553; UBI371; 1.
 CC K01554; UBI372; 1.
 CC K01555; UBI373; 1.
 CC K01556; UBI374; 1.
 CC K01557; UBI375; 1.
 CC K01558; UBI376; 1.
 CC K01559; UBI377; 1.
 CC K01560; UBI378; 1.
 CC K01561; UBI379; 1.
 CC K01562; UBI380; 1.
 CC K01563; UBI381; 1.
 CC K01564; UBI382; 1.
 CC K01565; UBI383; 1.
 CC K01566; UBI384; 1.
 CC K01567; UBI385; 1.
 CC K01568; UBI386; 1.
 CC K01569; UBI387; 1.
 CC K01570; UBI388; 1.
 CC K01571; UBI389; 1.
 CC K01572; UBI390; 1.
 CC K01573; UBI391; 1.
 CC K01574; UBI392; 1.
 CC K01575; UBI393; 1.
 CC K01576; UBI394; 1.
 CC K01577; UBI395; 1.
 CC K01578; UBI396; 1.
 CC K01579; UBI397; 1.
 CC K01580; UBI398; 1.
 CC K01581; UBI399; 1.
 CC K01582; UBI400; 1.
 CC K01583; UBI401; 1.
 CC K01584; UBI402; 1.
 CC K01585; UBI403; 1.
 CC K01586; UBI404; 1.
 CC K01587; UBI405; 1.
 CC K01588; UBI406; 1.
 CC K01589; UBI407; 1.
 CC K01590; UBI408; 1.
 CC K01591; UBI409; 1.
 CC K01592; UBI410; 1.
 CC K01593; UBI411; 1.
 CC K01594; UBI412; 1.
 CC K01595; UBI413; 1.
 CC K01596; UBI414; 1.
 CC K01597; UBI415; 1.
 CC K01598; UBI416; 1.
 CC K01599; UBI417; 1.
 CC K01600; UBI418; 1.
 CC K01601; UBI419; 1.
 CC K01602; UBI420; 1.
 CC K01603; UBI421; 1.
 CC K01604; UBI422; 1.
 CC K01605; UBI423; 1.
 CC K01606; UBI424; 1.
 CC K01607; UBI425; 1.
 CC K01608; UBI426; 1.
 CC K01609; UBI427; 1.
 CC K01610; UBI428; 1.
 CC K01611; UBI429; 1.
 CC K01612; UBI430; 1.
 CC K01613; UBI431; 1.
 CC K01614; UBI432; 1.
 CC K01615; UBI433; 1.
 CC K01616; UBI434; 1.
 CC K01617; UBI435; 1.
 CC K01618; UBI436; 1.
 CC K01619; UBI437; 1.
 CC K01620; UBI438; 1.
 CC K01621; UBI439; 1.
 CC K01622; UBI440; 1.
 CC K01623; UBI441; 1.
 CC K01624; UBI442; 1.
 CC K01625; UBI443; 1.
 CC K01626; UBI444; 1.
 CC K01627; UBI445; 1.
 CC K01628; UBI446; 1.
 CC K01629; UBI447; 1.
 CC K01630; UBI448; 1.
 CC K01631; UBI449; 1.
 CC K01632; UBI450; 1.
 CC K01633; UBI451; 1.
 CC K01634; UBI452; 1.
 CC K01635; UBI453; 1.
 CC K01636; UBI454; 1.
 CC K01637; UBI455; 1.
 CC K01638; UBI456; 1.
 CC K01639; UBI457; 1.
 CC K01640; UBI458; 1.
 CC K01641; UBI459; 1.
 CC K01642; UBI460; 1.
 CC K01643; UBI461; 1.
 CC K01644; UBI462; 1.
 CC K01645; UBI463; 1.
 CC K01646; UBI464; 1.
 CC K01647; UBI465; 1.
 CC K01648; UBI466; 1.
 CC K01649; UBI467; 1.
 CC K01650; UBI468; 1.
 CC K01651; UBI469; 1.
 CC K01652; UBI470; 1.
 CC K01653; UBI471; 1.
 CC K01654; UBI472; 1.
 CC K01655; UBI473; 1.
 CC K01656; UBI474; 1.
 CC K01657; UBI475; 1.
 CC K01658; UBI476; 1.
 CC K01659; UBI477; 1.
 CC K01660; UBI478; 1.
 CC K01661; UBI479; 1.
 CC K01662; UBI480; 1.
 CC K01663; UBI481; 1.
 CC K01664; UBI482; 1.
 CC K01665; UBI483; 1.
 CC K01666; UBI484; 1.
 CC K01667; UBI485; 1.
 CC K01668; UBI486; 1.
 CC K01669; UBI487; 1.
 CC K01670; UBI488; 1.
 CC K01671; UBI489; 1.
 CC K01672; UBI490; 1.
 CC K01673; UBI491; 1.
 CC K01674; UBI492; 1.
 CC K01675; UBI493; 1.
 CC K01676; UBI494; 1.
 CC K01677; UBI495; 1.
 CC K01678; UBI496; 1.
 CC K01679; UBI497; 1.
 CC K01680; UBI498; 1.
 CC K01681; UBI499; 1.
 CC K01682; UBI500; 1.
 CC K01683; UBI501; 1.
 CC K01684; UBI502; 1.
 CC K01685; UBI503; 1.
 CC K01686; UBI504; 1.
 CC K01687; UBI505; 1.
 CC K01688; UBI506; 1.
 CC K01689; UBI507; 1.
 CC K01690; UBI508; 1.
 CC K01691; UBI509; 1.
 CC K01692; UBI510; 1.
 CC K01693; UBI511; 1.
 CC K01694; UBI512; 1.
 CC K01695; UBI513; 1.
 CC K01696; UBI514; 1.
 CC K01697; UBI515; 1.
 CC K01698; UBI516; 1.
 CC K01699; UBI517; 1.
 CC K01700; UBI518; 1.
 CC K01701; UBI519; 1.
 CC K01702; UBI520; 1.
 CC K01703; UBI521; 1.
 CC K01704; UBI522; 1.
 CC K01705; UBI523; 1.
 CC K01706; UBI524; 1.
 CC K01707; UBI525; 1.
 CC K01708; UBI526; 1.
 CC K01709; UBI527; 1.
 CC K01710; UBI528; 1.
 CC K01711; UBI529; 1.
 CC K01712; UBI530; 1.
 CC K01713; UBI531; 1.
 CC K01714; UBI532; 1.
 CC K01715; UBI533; 1.
 CC K01716; UBI534; 1.
 CC K01717; UBI535; 1.
 CC K01718; UBI536; 1.
 CC K01719; UBI537; 1.
 CC K01720; UBI538; 1.
 CC K01721; UBI539; 1.
 CC K01722; UBI540; 1.
 CC K01723; UBI541; 1.
 CC K01724; UBI542; 1.
 CC K01725; UBI543; 1.
 CC K01726; UBI544; 1.
 CC K01727; UBI545; 1.
 CC K01728; UBI546; 1.
 CC K01729; UBI547; 1.
 CC K01730; UBI548; 1.
 CC K01731; UBI549; 1.
 CC K01732; UBI550; 1.
 CC K01733; UBI551; 1.
 CC K01734; UBI552; 1.
 CC K01735; UBI553; 1.
 CC K01736; UBI554; 1.
 CC K01737; UBI555; 1.
 CC K01738; UBI556; 1.
 CC K01739; UBI557;


```

DB 311 TVFRTSLRGKLVYRYSVQFLEK 332
|| : : | : : ||
[6]
RN SEQUENCE FROM N.A.
RP SPECIES=E. flexneri, STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
EX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.;
RL Infect. Immun. 71:2775-2786(2003).
RN [7]
RN CHARACTERIZATION, AND MUTANT UBIE401.
RP SPECIES=E. coli;
RC MEDLINE=97197541; PubMed=9045837;
EX Lee P.T., Hau A.Y., Ha H.T., Clarke C.F.;
RA "A C-methyltransferase involved in both ubiquinone and menaquinone
RT biosynthesis: isolation and identification of the Escherichia coli
RT ubiE gene."
RL J. Bacteriol. 179:1748-1754(1997).
RN [8]
RN FUNCTION: Methyltransferase required for the conversion of
RC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the
CC conversion of 2-polyphenyl-6-methoxy-1,4-benzoquinol (DMQKH2) to
CC 2-polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQKH2).
RN [9]
RN CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyphenyl-6-
CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
CC polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol.
RN [10]
RN CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
CC S-adenosyl-L-homocysteine + menaquinol.
RN [11]
RN PATHWAY: Menaquinone biosynthesis; last step.
RN [12]
RN PATHWAY: Ubiquinone biosynthesis.
RN [13]
RN SIMILARITY: Belongs to the ubiE family.
RN [14]
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
RN [15]
RN EMBL; M67049; AAC7628.1; -
DR EMBL; AS000459; AAC76836.1; -
DR EMBL; AS005614; AAGS9029.1; -
DR EMBL; AF002567; BAB38186.1; -
DR EMBL; AB015398; AAN45346.1; -
DR EMBL; AB016990; AAP18852.1; -
DR PIR; A86071; A86071.
DR PIR; B65188; B65188.
DR PIR; C98224; C98224.
DR ECGene; B011473; ubiE.
DR HAMAP; MF_01813; -; 1.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bind.
DR InterPro; IPR004034; Ubi/men Mentransf.
DR InterPro; IPR004033; UbiE/COQ5 Mentrif.
DR Pfam; PF01209; UbiE methyltrans; 1.
DR PROSITE; PS01183; UBI1; 1.
DR PROSITE; PS01184; UBI2; 1.
DR PROSITE; PS01185; UBI3; 1.
DR Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
DR Methyltransferase; Complete proteome.
FT MUTAGEN 142 142 G->D: IN UBIE401; DEFECTIVE.
FT CONFLICT 101 101 L -> P (IN REF. 2).
SQ SEQUENCE 251 AA; 28073 MW; B1EE209BDC68F5E CRC64;
Query Match 6.3%; Score 158.5; DB 1; Length 251;
Best Local Similarity 28.0%; Pred. NO. 8e-05;
Matches 60; Conservative 39; Mismatches 92; Indels 23; Gaps 10;
OY 245. VSTGGYETTKER-VSMLLDKPQKVLGVGCGIGG--GDFYMAETFDVEVVGPDLSVNMIS 301
DB 43 MSFGIHLRWKRFIDCSGVRGQTVLDLGGTGTAKFSRLVGTCKVLLADINESMLX 102

```

QY 302 FALF--RSLGKCAVEFEVADCTKINYPDNSFDVYSDTILHIQDKPALFSPYKWLKP 359
 DB 103 MREKRLNIGVGNVEYVQANAEALPPFDNFPCITISFGLRNVTDKOKALSMTRVLKP 162
 QY 360 GSKVLISDYCKKAGPPSPFAAYIKQBGYDLHDVKEYGQML-KDAGFVDVLAEDTEQFI 418
 DB 163 GGRLLVLEFSKPIIEPLSK--AY---DAYSFHLVPRIGSLVANDADSYRYLAES-----I 212
 QY 419 RVLRKELETVRKEKDVFISSD--FSEEDYNDIVGG 450
 DB 213 R-MHPDQDTLK-----ANMQDAGFSDVYNYLTAG 241

RESULT 14

UBIE_VIBU STANDARD; PRT; 260 AA.
 AC QSDDEP9;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE
 DE (EC 2.1.1.-)
 GN UBIE OR V10909
 OS Vibrio vulnificus
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.B.;
 RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";
 RT Submitted (DEC 2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Methyltransferase required for the conversion of
 CC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the
 CC conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to
 CC 2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-
 CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
 CC polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
 CC S-adenosyl-L-homocysteine + menaquinol.
 CC -!- PATHWAY: Menaquinone biosynthesis; last step.
 CC -!- PATHWAY: Ubiquinone biosynthesis.
 CC -!- SIMILARITY: Belongs to the ubiE family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AS016800; AAC09411.1;
 DR HAMAP; MF_01813; -;
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 DR InterPro; IPR004034; Ubi/men.Metransf.
 DR Pfam; PF01209; UbiE_methyltran.1.
 DR PROSITE; PS01183; UBIE_1; 1.
 DR PROSITE; PS01184; UBIE_2; 1.
 DR Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
 DR Methyltransferase; Complete proteome.
 SQ SEQUENCE 260 AA; 28971 MW; 813262BDF89BE802 CRC64;

Query Match 63% Score: 158.5; DB 1; Length 260;
 Best Local Similarity 26.1%; Pred. No. 8.4e-05;
 Matches 57; Conservative 45; Mismatches 85; Indels 31; Gaps 11;

QY 245 VSTGYETTKEF--VSMJDLKPGQKVLVVGCGIGGGDFYMAETFDVEVVG-----FDLS 296
 DB 52 MSGGTHRLWKEFTDCSGARPGQVLDLGGGTG-----LTAKFS-RIVGKGVILADIN 106
 QY 297 VMISFALE--RSLGKCAVEFEVADCTKINYPDNSFDVYSDTILHIQDKPALFSPFY 354
 DB 107 NSMLNVRGDKLRDSGIVGNVHYVQANAEELPPDDYDFDIITISFCLRNVTDKOKALRSMF 166
 QY 355 KWLKGGKVLISDYCKKA-GPPSPFAAYIKQBGYDLHDVKEYGQML-KDAGFVDVLAED 412
 DB 167 RVLRKELETVRKEKDVFISSD-----FSEEDYNDIVGG 450
 QY 413 RTEQFIRVLRKELETVRKEKDVFISSDSEEDYNDIVGG 450
 DB 221 -----IR-MHPDQDTLKGMQE--AGFENTSYNYLTGG 250

RESULT 15

UBIE_NEIMA STANDARD; PRT; 245 AA.
 AC Q8JVE3;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquinone/menaquinone biosynthesis methyltransferase ubiE
 DE (EC 2.1.1.-)
 GN UBIE OR NMA0956.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=22491 / Serogroup A / Serotype 4A;
 RX MEDLINE=2022556; PubMed=10761919;
 RA Parkhill J., Achtman G., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli M., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
 RA Jagals K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RA "Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491.";
 RL Nature 404:502-506(2000).
 CC -!- FUNCTION: Methyltransferase required for the conversion of
 CC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) and the
 CC conversion of 2-polyprenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to
 CC 2-polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol (DMQH2) (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-
 CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
 CC polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
 CC S-adenosyl-L-homocysteine + menaquinol.
 CC -!- PATHWAY: Menaquinone biosynthesis; last step.
 CC -!- PATHWAY: Ubiquinone biosynthesis.
 CC -!- SIMILARITY: Belongs to the ubiE family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AL162754; CAB84226.1; -;
 DR FIR; C81942; C81942.
 DR HAMAP; MF_01813; -;
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 DR InterPro; IPR004034; Ubi/men.Metransf.
 DR InterPro; IPR004033; UbiE/COQ5_Metrif.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 26, 2004, 13:11:41 ; Search time 41 Seconds
(without alignments)
3640.003 Million cell updates/sec

Title: US-10-031-331B-40

Perfect score: 2515

Sequence: 1 HTVDLTIEMMLDSQASDLD.....KLRRATGQRWGLPVAKK 473

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2515	100.0	494	10 Q85287	Q85287 suada Japo
2	2063	82.0	491	10 Q9AXH3	Q9AXH3 lycopersico
3	2044.5	81.3	493	10 Q84SA4	Q84SA4 aster tripo
4	2009	79.9	491	10 Q7XJZ2	Q7XJZ2 brassica na
5	1968	78.3	499	10 Q8LJ10	Q8LJ10 oryza sativ
6	1918.5	76.3	498	10 Q8VVK1	Q8VVK1 triticum ae
7	1515	60.2	376	10 Q8L7A8	Q8L7A8 arabidopsis
8	516.5	20.5	437	5 Q22933	Q22933 caenorhabdi
9	476	18.9	266	5 Q81DQ9	Q81DQ9 plasmodium
10	433	16.4	460	5 Q86NE3	Q86NE3 caenorhabdi
11	413	16.4	475	5 Q23552	Q23552 caenorhabdi
12	413	16.4	484	5 Q95PW7	Q95PW7 caenorhabdi
13	288.5	11.5	264	16 Q98VFB	Q98VFB rhizobium 1
14	238	9.5	351	5 Q8IFX3	Q8IFX3 caenorhabdi
15	229.5	9.1	270	2 Q83X76	Q83X76 streptomyc
16	225.5	9.0	276	2 Q846Y0	Q846Y0 streptomyc

17	222	8.8	565	2	Q9KJ20	Q9KJ20 actinopolya
18	218.5	8.7	283	2	Q9X5Q9	Q9X5Q9 streptomyc
19	217.5	8.6	363	10	Q41587	Q41587 triticum ae
20	217.5	8.6	363	10	Q41586	Q41586 triticum ae
21	214	8.5	280	16	Q81W13	Q81W13 anabaena sp
22	211.5	8.4	346	10	Q82720	Q82720 nicotiana t
23	210.5	8.4	317	16	Q55809	Q55809 synecocyst
24	209.5	8.3	272	2	Q8K152	Q8K152 amocollatops
25	206.5	8.2	273	2	Q8K152	Q8K152 amocollatops
26	206.5	8.2	283	2	Q8K294	Q8K294 nocardia ae
27	203	8.1	290	16	Q8KBN7	Q8KBN7 chlorobium
28	200.5	8.0	279	2	Q9KJ21	Q9KJ21 ecotiorho
29	200	8.0	278	2	Q9EVI2	Q9EVI2 streptomyc
30	199.5	7.9	349	10	Q82434	Q82434 nicotiana t
31	198	7.9	344	10	Q49215	Q49215 zea mays (m
32	198	7.9	344	10	P93852	P93852 zea mays (m
33	197	7.8	276	2	Q83WG2	Q83WG2 streptomyc
34	196	7.8	330	10	Q84M50	Q84M50 oryza sativ
35	194.5	7.7	283	2	Q9ALN3	Q9ALN3 saccharopol
36	193	7.7	367	10	Q43445	Q43445 glycine max
37	190	7.6	349	10	Q82426	Q82426 oryza sativ
38	189.5	7.5	280	2	Q83WF7	Q83WF7 streptomyc
39	187.5	7.5	330	16	Q8YV60	Q8YV60 anabaena sp
40	186.5	7.4	336	10	Q9LM02	Q9LM02 arabidopsis
41	185.5	7.4	227	17	Q57965	Q57965 pyrococcus
42	184.5	7.3	318	16	P74388	P74388 synecocyst
43	184.5	7.3	336	10	Q8LKW1	Q8LKW1 arabidopsis
44	184	7.3	277	2	Q83WC3	Q83WC3 aphanotheca
45	183.5	7.3	387	16	Q9PNB2	Q9PNB2 campylobact

ALIGNMENTS

RESULT 1

Q85287	PRELIMINARY; PRT; 494 AA.
ID	Q85287
AC	Q85287; TREMBLrel. 24, Created)
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT	01-JUN-2003 (TREMBLrel. 25, Last annotation update)
DE	Phosphoethanolamine N-methyltransferase.
GN	PEAMT.
OS	Suadea japonica.
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Caryophyllales; Amaranthaceae; Suadea.
OX	NCBI_TaxID=90346;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Yamada A., Nozawa G.T., Tanimoto S., Ozeki Y.;
RT	"Glycinebetaine synthesis in Suadea japonica."
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB080186; BAC57432.1;
DR	GO; GO:0008755; F:S-adenosylmethionine-dependent methyltransf. . . ; IEA.
DR	GO; GO:0016740; F:transferase activity; IEA.
DR	InterPro; IPR001601; Methyltransf.
DR	InterPro; IPR000051; SAM bind.
DR	InterPro; IPR004033; UbiE/COO5_Metrf.
DR	Pfam; PF01209; Ubie_methyltran; 1.
KW	Transferase; Methyltransferase.
SQ	SEQUENCE 494 AA; 56562 MW; BC613P9097BD3AE3 CRC64;

Query Match 100.0%; Score 2515; DB 10; Length 494;
Best Local Similarity 100.0%; Pred. No. 1.9e-188;
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HTVDLTIEMMLDSQASDLDKEPERPELMLPPEKCLLELCAGIGRTGSLAKAQV 60

22 HTVDLTIEMMLDSQASDLDKEPERPELMLPPEKCLLELCAGIGRTGSLAKAQV 81

Qy 61 IALDPIESALIKNVKNGKVKFCADVTSTLSFPFHSLDVTFESWMLLYLSERVE 120

Db 82 IALDPIESAIKNEVINGHYKNVCMKADVTSPITLSPFHSLDVIIPSNMLLWYLSDEVE 141
QY 121 NLVERMLKWLKPGGYIPFRESCHQGDHRSNTHYREPRFTKPEKCHLODGSNS 180
Db 142 NLVERMLKWLKPGGYIPFRESCHQGDHRSNTHYREPRFTKPEKCHLODGSNS 201
QY 181 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQRPFLDTSQYKNSILRYERV 240
Db 202 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQRPFLDTSQYKNSILRYERV 261
QY 241 GPGVSTGGYETTKFVSMMLDLKPGQKVDVGGCIGGGDFYMAETFDVVEVGGFSLVNM 300
Db 262 GPGVSTGGYETTKFVSMMLDLKPGQKVDVGGCIGGGDFYMAETFDVVEVGGFSLVNM 321
QY 301 SPALERSIGLKAVERFVADCTKINYPDINSFDVYSRDTILHIDKPALEFSPKMLKPG 360
Db 322 SPALERSIGLKAVERFVADCTKINYPDINSFDVYSRDTILHIDKPALEFSPKMLKPG 381
QY 361 GKVLISDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 420
Db 382 GKVLISDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 441
QY 421 LRKELETVEKOVPIISDFSEEDYNDIVGWNDKLRTAKGEORWGLFVAKKK 473
Db 442 LRKELETVEKOVPIISDFSEEDYNDIVGWNDKLRTAKGEORWGLFVAKKK 494

RESULT 2
Q9AMX3 PRELIMINARY; PRT; 491 AA.
ID Q9AMX3
AC Q9AMX3; 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Phosphoethanolamine N-methyltransferase.
OS Lycopersicon esculentum (tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Solanum.
CX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Parani M., Parida A.;
RT Characterization of a cDNA for phosphoethanolamine N-methyltransferase.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228858; AAG59894.1; -.
DR GO; GO:0005344; F:oxyanion transporter activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006810; P:transferase activity; IEA.
DR InterPro; IPR000896; Hemocyanin.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bind.
DR PROSITE; PS00210; HEMOCYANIN 2; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 491 AA; 55947 MW; D3GSCDD8512733E7 CRC64;

Query Match 82.0%; Score 2063; DB 10; Length 491;
Best Local Similarity 80.1%; Pred. No. 4.7e-153;
Matches 379; Conservative 46; Mismatches 148; Indels 0; Gaps 0;

QY 1 HTVDLTTEAMMLDSQASDLDEKPEPILSMPLPPEGLKLELGLAGIGRFTGELAKAGQV 60
Db 19 HTAELTYEAMMLDSKAADLDEKPEPILSMPLPPEGLKLELGLAGIGRFTGELAKAGQV 78
QY 61 IALDPIESAIKNEVINGHYKNVCMKADVTSPITLSPFHSLDVIIPSNMLLWYLSDEVE 120
Db 79 IALDPIESAIKNEVINGHYKNVCMKADVTSPITLSPFHSLDVIIPSNMLLWYLSDEVE 138
QY 121 NLVERMLKWLKPGGYIPFRESCHQGDHRSNTHYREPRFTKPEKCHLODGSNS 180
Db 139 ALVERWMLKVGCGHIFPESCFHQSGDHKRNTHYREPRFTKPEKCHLODGSNS 198

QY 181 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQRPFLDTSQYKNSILRYERV 240
Db 199 FELSLAGCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQRPFLDTSQYKNSILRYERV 258
QY 241 GPGVSTGGYETTKFVSMMLDLKPGQKVDVGGCIGGGDFYMAETFDVVEVGGFSLVNM 300
Db 259 GPGVSTGGYETTKFVSMMLDLKPGQKVDVGGCIGGGDFYMAETFDVVEVGGFSLVNM 318
QY 301 SPALERSIGLKAVERFVADCTKINYPDINSFDVYSRDTILHIDKPALEFSPKMLKPG 360
Db 319 SPALERSIGLKAVERFVADCTKINYPDINSFDVYSRDTILHIDKPALEFSPKMLKPG 378
QY 361 GKVLISDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 420
Db 379 GKVLISDYCKKAGPPSPFAAYIKQGYDLHDVKEYGQMLKADGFDVLAEDRTQFIRV 438
QY 421 LRKELETVEKOVPIISDFSEEDYNDIVGWNDKLRTAKGEORWGLFVAKKK 473
Db 439 LRKELETVEKOVPIISDFSEEDYNDIVGWNDKLRTAKGEORWGLFVAKKK 491

RESULT 3
Q84SA4 PRELIMINARY; PRT; 493 AA.
ID Q84SA4
AC Q84SA4; 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 25, Last annotation update)
DE Phosphoethanolamine N-methyltransferase.
OS Aster tripolium (sea aster).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Asteraceae; Aster.
CX NCBI_TaxID=74787;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeda M., Uno Y., Kanechi M., Inagaki N.;
RT "Analysis of nine cDNAs for salt-inducible genes in the halophyte sea aster.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB090883; BAC57960.1; -.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bind.
KW Methyltransferase; Transferase.
SQ SEQUENCE 493 AA; 56002 MW; 9F2C7369192B6DA5 CRC64;

Query Match 81.3%; Score 2044.5; DB 10; Length 493;
Best Local Similarity 78.9%; Pred. No. 1.3e-151;
Matches 374; Conservative 51; Mismatches 48; Indels 1; Gaps 1;

QY 1 HTVDLTTEAMMLDSQASDLDEKPEPILSMPLPPEGLKLELGLAGIGRFTGELAKAGQV 60
Db 20 HSDVLTVESMLDSMASDLDEKPEPILSMPLPPEGLKLELGLAGIGRFTGELAKAGQV 79
QY 61 IALDPIESAIKNEVINGHYKNVCMKADVTSPITLSPFHSLDVIIPSNMLLWYLSDEVE 120
Db 80 IALDPIESAIKNEVINGHYKNVCMKADVTSPITLSPFHSLDVIIPSNMLLWYLSDEVE 139
QY 121 NLVERMLKWLKPGGYIPFRESCHQGDHRSNTHYREPRFTKPEKCHLODGSNS 180
Db 140 DIABRFLKWLKVGCGHIFPESCFHQSGDHKRNTHYREPRFTKPEKCHLODGSNS 199
QY 181 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQRPFLDTSQYKNSILRYERV 239
Db 200 YELSLSCCKICIGAYVRNKKQNIQISWLQKVDSDKDKGQRPFLDTSQYKNSILRYERV 259
QY 240 GPGVSTGGYETTKFVSMMLDLKPGQKVDVGGCIGGGDFYMAETFDVVEVGGFSLVNM 299
Db 260 GPGVSTGGYETTKFVSMMLDLKPGQKVDVGGCIGGGDFYMAETFDVVEVGGFSLVNM 319

```

QY 300 ISFALERSIGLKCAVEFEVADCTKINYPDNSFDVYISRDITLHIQDKPALERSFYKMLKP 359
Db 320 IAFALERSIGLQCSVEFEVADCTKSPDNSFDVYISRDITLHIQDKPALFRITFKMLKP 379
QY 360 GKVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 419
Db 380 GKVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 439
QY 420 VLKLELTVKEKDVIFSDSEEDYNDIVGWNDRKLRATKGEQRMGLFVAKK 473
Db 440 VLKLELTVKEKDVIFSDSEEDYNDIVGWNDRKLRATKGEQRMGLFVAKK 493

RESULT 4
Q7XJ22 PRELIMINARY; PRT; 491 AA.
ID Q7XJ22
AC Q7XJ22
DT 01-OCT-2003 (TREMBlrel. 25, Last Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE Phosphothanolamine N-methyltransferase.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euraoidii; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye C., Li J., Yang J., Wang B.;
RT "Plant gene cloning and expression analysis under different
RT stresses";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RW EMBL: AY319479; AAP3582.1;
KW Methyltransferase; Transferase.
SQ SEQUENCE 491 AA; 55398 MW; A08C3318737031EE CRC64;

```

```

Query Match 79.9%; Score 2009; DB 10; Length 491;
Best Local Similarity 78.0%; Pred. No. 7.9e-149;
Matches 368; Conservative 52; Mismatches 52; Indels 0; Gaps 0;

QY 1 HTVDITIAMLDLSDQASDLKDEERPEILSMPLPEGLKLELGAGIGRTGELAKAGOV 60
Db 19 HSADLTVEANMLDSRAADLDEERPEILSLPPYEGKSVLELGAGIGRTGELAKAGEL 78
QY 61 IALDFIESAIKKNEVINGHYKNVCFMCAVTSPTLSPPPHSLDVTSPNNMLVLSDEIVE 120
Db 79 IALDFISVINKNESVINGHYKNVCFMCAVTSPTLSPPPHSLDVTSPNNMLVLSDEIVE 138
QY 121 NLVEMRLKWLKPGYIFPSCFQSGDHKRNKSNPTHYREPRPYTKAPKECHLODGSNGS 180
Db 139 LLVEMVGVWIKVGGYIFPSCFQSGDHKRNKSNPTHYREPRPYTKAPKECHLODGSNGS 198
QY 181 YELSLSCCKICGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 240
Db 199 FELSNIGKCCIGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 258
QY 241 GPGYSTGGYETTKFVSMGLDKPGQKLVGGCGIGGDFYMAETPDVSVVGFDSLNNMI 300
Db 259 GPGYSTGGYETTKFVSMGLDKPGQKLVGGCGIGGDFYMAETPDVSVVGFDSLNNMI 318
QY 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVYISRDITLHIQDKPALFRSFKMLKPG 360
Db 319 SPALERAIGLNCSEFEVADCTKHYKSPNSKDVYISRDITLHIQDKPALFRITFKMLKPG 378
QY 361 GKVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 420
Db 379 GKVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 438
QY 421 LRKLELTVKEKDVIFSDSEEDYNDIVGWNDRKLRATKGEQRMGLFVAKK 472
Db 439 LRKLELTVKEKDVIFSDSEEDYNDIVGWNDRKLRATKGEQRMGLFVAKK 490

```

```

RESULT 5
Q8LJ10 PRELIMINARY; PRT; 499 AA.
ID Q8LJ10
AC Q8LJ10
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Putative phosphothanolamine methyltransferase.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0431H09.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP003248; BAC10708.1;
DR Gramene; Q8LJ10;
DR GO: GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bind.
KW Transferase; Methyltransferase.
SQ SEQUENCE 499 AA; 56786 MW; ADDF3D04E62D18BE CRC64;

```

```

Query Match 78.3%; Score 1968; DB 10; Length 499;
Best Local Similarity 76.1%; Pred. No. 1.3e-145;
Matches 359; Conservative 52; Mismatches 61; Indels 0; Gaps 0;

QY 1 HTVDITIAMLDLSDQASDLKDEERPEILSMPLPEGLKLELGAGIGRTGELAKAGOV 60
Db 28 HSKDLTVEANMLDSRAADLDEERPEILSLPPYEGKSVLELGAGIGRTGELAKAGV 87
QY 61 IALDFIESAIKKNEVINGHYKNVCFMCAVTSPTLSPPPHSLDVTSPNNMLVLSDEIVE 120
Db 88 LAMDFIESVINKNESVINGHYKNVCFMCAVTSPTLSPPPHSLDVTSPNNMLVLSDEIVE 147
QY 121 NLVEMRLKWLKPGYIFPSCFQSGDHKRNKSNPTHYREPRPYTKAPKECHLODGSNGS 180
Db 148 KLVKRWMLKVGYYIFPSCFQSGDHKRNKSNPTHYREPRPYTKAPKECHLODGSNGS 207
QY 181 YELSLSCCKICGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 240
Db 208 FELSVLTCKCVGAYVRNKKQNOISWLWQVDSKDKGFORPLDTQSYKNSILYERVP 267
QY 241 GPGYSTGGYETTKFVSMGLDKPGQKLVGGCGIGGDFYMAETPDVSVVGFDSLNNMI 300
Db 268 GPGYSTGGYETTKFVSMGLDKPGQKLVGGCGIGGDFYMAETPDVSVVGFDSLNNMI 327
QY 301 SPALERSIGLKCAVEFEVADCTKINYPDNSFDVYISRDITLHIQDKPALFRSFKMLKPG 360
Db 328 SPALERAIGRCSFEFEVADCTKTKYDPNTDFVYISRDITLHIQDKPSLFSFKMLKPG 387
QY 361 GKVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 420
Db 388 GKVLSIDYCKKAGPPSPFAAYIKQGYDLHDVKEYGOMLKDAGFVDVLAEDRTQPIR 447
QY 421 LRKLELTVKEKDVIFSDSEEDYNDIVGWNDRKLRATKGEQRMGLFVAKK 472
Db 448 LRKLELTVKEKDVIFSDSEEDYNDIVGWNDRKLRATKGEQRMGLFVAKK 499

```

```

RESULT 6
Q8VYX1 PRELIMINARY; PRT; 496 AA.
ID Q8VYX1
AC Q8VYX1
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

```

01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Phosphoethanolamine methyltransferase.
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
 OC Triticaceae; Triticum.
 NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Brenneise Charon J.-B., Breton G., Danyluk J., Muzac I., Ibrahim R.,
 SA Sarhan P.;
 RT "Molecular and biochemical characterization of a cold regulated
 phosphoethanolamine methyltransferase from wheat."
 RL Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY065971; AAL0895.1;
 DR GO; GO:0008757; P:8-adenosylmethionine-dependent methyltransf. . . IEA.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 KW Transferase; Methyltransferase.
 SQ SEQUENCE 498 AA; 56858 MW; 37BB7134E2DEA148 CRC64;
 Query Match 76 34; Score 1918.5; DB 10; Length 498;
 Best Local Similarity 73.84; Pred. No. 9.9e-142;
 Matches 349; Conservative 59; Mismatches 64; Indels 1; Gaps 1;
 QY 1 HTVDTTAMMLDSQADLDKEERPEILSMPLPESKCLLELGAGIGRTGELAEKAGOV 60
 DB 26 HSKDLTVSMMLDSQADLDKEERPEILSMPLPESKCLLELGAGIGRTGELAEKAGOV 85
 QY 61 IALDPTESAIKNEVINGHYKVKFCADVTSPTLSFPFPHSLDVIFSNWLLMYLSDREV 119
 DB 86 IALDPTESAIKNEVINGHYKVKFCADVTSPTLSFPFPHSLDVIFSNWLLMYLSDREV 145
 QY 120 EMLVERMLKLGKGYIFPESCFHQSDHKNKSNPTHYEPFYTKAFKCHLQDGSN 179
 DB 146 EKLIGRVKMLPGHIFPESCFHQSDHKNKSNPTHYEPFYTKAFKCHLQDGSN 205
 QY 180 SYELSLSCCKTGAYVRKKNQNIQISMLWQVDSKDKGFORFLDTSQKNSILYERVF 239
 DB 206 SPFLSVTSKCIKAYVRKKNQNIQISMLWQVDSKDKGFORFLDTSQKNSILYERVF 265
 QY 240 FGKGYVSTGYETTKFVSMPLDKPGQKVLVGGGIGGDFYMAETFDVVEVGFDSVNM 299
 DB 266 FGKGYVSTGYETTKFVSMPLDKPGQKVLVGGGIGGDFYMAETFDVVEVGFDSVNM 325
 QY 300 ISFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDITLHIQDKPALFRSFKWLP 359
 DB 326 VSPFAIRAGRSVFEVADCTKINYPDSNDFVYISRDITLHIQDKPALFRSFKWLP 385
 QY 360 GGVKLVSDYCKKAGPSPPEFAAYIKQGYDLHDVKEYGQMLKQAGFVDVLAEDRTQFTR 419
 DB 386 GGVKLVSDYCKKAGPSPPEFAAYIKQGYDLHDVKEYGQMLKQAGFVDVLAEDRTQFTR 445
 QY 420 VLKKELETKEKQVPSPEPSRDYDVGWNNDKLRRTAKGKRGMLFVAKK 472
 DB 446 VLKKELETKEKQVPSPEPSRDYDVGWNNDKLRRTAKGKRGMLFVAKK 498
 RESULT 7
 Q8L7A8 PRELIMINARY; PRT; 376 AA.
 ID Q8L7A8
 AC Q8L7A8
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Phosphoethanolamine N-methyltransferase, putative.
 GN ATIG73600.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 NCBI_TaxID=3702;

SEQUENCE FROM N.A.
 RP Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
 RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
 RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
 RA Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Shimizu P., Tang C.C., Toroumi M., Wallander E.K., Wong C.,
 RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
 RA Theologis A., Davis R.W.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY136372; AA97038.1;
 DR GO; GO:0008757; P:8-adenosylmethionine-dependent methyltransf. . . IEA.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000051; SAM bind.
 KW Transferase; Methyltransferase.
 SQ SEQUENCE 376 AA; 42811 MW; 2B4F6BE2B112B3ED CRC64;
 Query Match 60.24; Score 1515; DB 10; Length 376;
 Best Local Similarity 81.14; Pred. No. 2.7e-110;
 Matches 279; Conservative 31; Mismatches 34; Indels 0; Gaps 0;
 QY 1 HTVDTTAMMLDSQADLDKEERPEILSMPLPESKCLLELGAGIGRTGELAEKAGOV 60
 DB 32 HSKDLTVSMMLDSQADLDKEERPEILSMPLPESKCLLELGAGIGRTGELAEKAGOV 91
 QY 61 IALDPTESAIKNEVINGHYKVKFCADVTSPTLSFPFPHSLDVIFSNWLLMYLSDREV 120
 DB 92 IALDPTESAIKNEVINGHYKVKFCADVTSPTLSFPFPHSLDVIFSNWLLMYLSDREV 151
 QY 121 NLVERMLKLGKGYIFPESCFHQSDHKNKSNPTHYEPFYTKAFKCHLQDGSN 180
 DB 152 DLAKGQVMTKGGIIFPESCFHQSDHKNKSNPTHYEPFYTKAFKCHLQDGSN 211
 QY 181 YELSLSCCKTGAYVRKKNQNIQISMLWQVDSKDKGFORFLDTSQKNSILYERVF 240
 DB 212 YELSLSCCKTGAYVRKKNQNIQISMLWQVDSKDKGFORFLDTSQKNSILYERVF 271
 QY 241 FGKGYVSTGYETTKFVSMPLDKPGQKVLVGGGIGGDFYMAETFDVVEVGFDSVNM 300
 DB 272 FGKGYVSTGYETTKFVSMPLDKPGQKVLVGGGIGGDFYMAETFDVVEVGFDSVNM 331
 QY 301 SFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDITLHIQ 344
 DB 332 SFALERSGLKCAVEFEVADCTKINYPDSNDFVYISRDITLHIQ 375
 RESULT 8
 Q22993 PRELIMINARY; PRT; 437 AA.
 ID Q22993
 AC Q22993
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN F54D11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Bristol N2;
 RX MEDLINE=99059613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Bristol N2;
 RA Pauley A., Gattung S.;
 RT "The sequence of C. elegans cosmid F54D11.";

```

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RT Waterston R.;
RA "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64834; AAB04824.1; -.
DR F1R; T29330; T29330.
DR WormPep; F54D11.1; CE11068.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.
DR InterPro; IPR001601; Methylbind.
DR InterPro; IPR000051; SAM bind.
DR Pfam; PF01209; UbiE/COG5 Metrf.
DR Hypothetical protein.
RW Hypothetical protein.
SQ SEQUENCE 437 AA; 49769 MW; 6675E262F67D08B CRC64;

Query Match 20.5%; Score 516.5; DB 5; Length 437;
Best Local Similarity 33.6%; Pred. No. 5.1e-32;
Matches 127; Conservative 66; Mismatches 164; Indels 21; Gaps 8;

QY 104 VIFSNLLMYLSDEE-VENLVERMLKWLKPGGVIFPFBSCFHQSDHKEKSNPTHYRPR 162
DB 68 LIFNALSQITNADLTDFPKATNATAGGVIRED-LKDCSDKQVRLTY---- 122

QY 163 FYTKAFKECHLDQSGNSYELSLSCKICIGAYVRNKKONQISLW----QKVDKDDKG- 218
DB 123 -----FDVFTTDSGNTGLDLYTDQVEH--SNTVEQNFLDFIFVFRKKVFAPTTDTAT 175

QY 219 --FQRLDTSQYKCNILRVERVFGPGVSTGYETTKFVSM-LDKPGQKVLVDVGGCI 275
DB 176 ITFDFDLKDTYWTGIDAVEMPGVNFISPGGYDENLKIIRPGDFKGGTMDLIGVGI 235

QY 276 GGGDFYKARTDVEVGFDSLVNMIISPALESGIKL-CAVEFEVADCTKINYPNSPDVI 334
DB 236 GGGARQVADFGVHVHGLDLSNMLALERLHEBKDSRKYSITDALVQPEDNSPDYV 295

QY 335 YSRDTILHIDKPAFLPSFYKWLKPGGKVLISDYCKAGPPSPFAAYIKQGYDLHDVK 394
DB 296 FSRDCIQHIDPTEKIFSLRIYKALKPGGKVLITMYGKGYGEQSDKFTYVAGRAFLKVLK 355

QY 395 EYGMQLKDGAFVDVLAERTQFIRVLKKELETVEKEDVFTSDSEEDYNDIVGWNKD 454
DB 356 EIADIANKTGVVNTENMPFKEILLEGHLEQNEAEPMFKTQREDSLISGWTDX 415

QY 455 IRTAKGQRGLFVAKK 472
DB 416 LGYIEKDNHNNFFLAQK 433

RESULT 9
ID Q81DQ9 PRELIMINARY; PRT; 266 AA.
AC Q81DQ9;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Phosphoethanolamine N-methyltransferase, putative (SC 2.1.1.103).
GN MAL13P1.214.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]_TaxID=36329;
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Brimman M., Fain A., Hall N., Aklin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL844509; -CAD52560.1; -.
DR GO; GO:000234; F:phosphoethanolamine N-methyltransferase act. . .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001601; Methyltransf.

Query Match 16.4%; Score 413; DB 5; Length 460;
Best Local Similarity 28.6%; Pred. No. 7e-24;
Matches 128; Conservative 70; Mismatches 202; Indels 48; Gaps 14;

DR InterPro; IPR000051; SAM bind.
RW Transferase; Methyltransferase.
SQ SEQUENCE 266 AA; 31043 MW; A7CBE98D882DBF74 CRC64;

Query Match 18.9%; Score 476; DB 5; Length 266;
Best Local Similarity 38.2%; Pred. No. 3.9e-29;
Matches 97; Conservative 54; Mismatches 99; Indels 4; Gaps 3;

QY 222 FLDTSQYKCNILRVERVFGPGVSTGYETTKFVSM-LDKPGGKVLVDVGGIGGGDFY 281
DB 13 FLENNQYTDGKVKYVFIFGENYISSGLEAYTKILSDIELANENSKVLDSGLGGCMY 72

QY 282 MAETFOVEVGFDSLVNMIISPALESGIKLCAVEFEVADCTKINYPNSPDVYISRDIL 341
DB 73 INEKYGANTHGDICSNIVNMANERVSG-NNKTIIPANDILTKFEPNNFDLIYSRDAIL 131

QY 342 H--IQDKPALPSFYKWLKPGGKVLISDYCKAGPP-SPEFAAYIKQGYDLHDVKYQG 398
DB 132 HLSLEKNKILFQKCYKWLKPTGTLITDYCAEKENWDDFEKYYVKQKYTLITVRYAD 191

QY 399 MKDAGFVDVLAERTQFIRVLKKELETVEKEDVFTSDSEEDYNDIVGWNKDART 458
DB 192 ILTACNFRNVSKDLSYNNQLLEVEHKYLNKESFEKLFSEKKEFISLDDGSKIKDS 251

QY 459 AKGBORGLFVAKK 472
DB 252 KRQQRNGYFRATX 265

RESULT 10
ID Q86NB3 PRELIMINARY; PRT; 460 AA.
AC Q86NB3;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein ZK622.3d.
GN ZK622.3.
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_TaxID=6239;
RP SEQUENCE FROM N.A.
RA Waterston R.;
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Leimbach D.;
RT "The sequence of C. elegans cosmid ZK622."
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; U39998; AAO38584.1; -.
DR WormPep; ZK622.3d; CE33505.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase. . .; IEA.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM bind.
RW Hypothetical protein.
SQ SEQUENCE 460 AA; 53540 MW; 74FF30F3E13CBBD7 CRC64;

Query Match 16.4%; Score 413; DB 5; Length 460;
Best Local Similarity 28.6%; Pred. No. 7e-24;
Matches 128; Conservative 70; Mismatches 202; Indels 48; Gaps 14;
```


QY 9 AMWDSQASDLKEERPEILSMPLPCKGKLELGGAGIGRETGLAKAGAGVIALDFIBS 68
 DB 24 SMWLNHSAELESSESSADILASLPLLNKQVVDIGAGIGRTTGLAETARWVLSDFIDS 83
 QY 69 AIKQNEVINGHYKQVPMKADVTSPFLSPHSLDVFPSNMLMYLSDEVENLVERMLX 128
 DB 84 FIKKQERNNAHLNINQVGDVAG--LKMSNSVDLVFTNMLMYLSDEETVEFIFNCR 141
 QY 129 WLKPGGYIFFRESCFQHS--GDHKKRS-----NPHYREPRFYTKAFKCHLQDGSNS 180
 DB 142 WLRSHGIVHLRESCEPSTGRSKAKSMHDTANANPHYRFSLSYLINLRAIRYRDVDNKL 201
 QY 181 YELSLSCCKIGAYVRNKNQNOISWLMQVSKDKGQFQFLDTSOYKNSILRYERVF 240
 DB 202 WRFNVQWSCSVPTTYIKRSNNRWQVHWAELKPAEDGAKGTSENLVELIKNTWQNEQAW 261
 QY 241 GPGYVSTGGYETTKFVSMMLDKPGQ-----KVLVGGCIGGGDFYMAETDVEVVG 292
 DB 262 DAKLDEKYYVTDKVFSALTSLSNSSTFLYTPRTVSPYCHINA--HTLAETFNANVMN 319
 QY 293 FDLNWMISFALERSIGLK-CAVEF-----EVAD-CTKINYPDNPSPDVIYSDTLIHQDK 346
 DB 320 TEIIPYRTSLTKSNLKDQVRFGNQSLTDSVTYMQKDALFDVVFATEFLSTVDDE 379
 QY 347 PALFRSPFYKWKPGCKVLISDYCKKAGPPSPFAAYIKORGVDLHDVKEYGOMLKDAGFV 406
 DB 380 --TIRQLPNVMSDGAKEPITL-----PVDEVNEAEMKOR-----IQBLGTYLK--SPT 423
 QY 407 DVLAE--DRTEQPIRLVLRKELETVEKEK 432
 DB 424 DVTDOCTEAOQRYE----KDHEQLRDEK 447

RESULT 11
 Q23552 ID Q23552 PRELIMINARY; PRT; 475 AA.
 AC Q23552;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein ZK622.3.
 GN ZK622.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Wilson R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Leimbach D.;
 RT "The sequence of C. elegans cosmid ZK622.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U39998; AAC038583.1; -
 DR PIR; T27936; T27936.
 DR WormPep; ZK622.3a; CE33504.
 DR GO; GO:008757; F1S-adenosylmethionine-dependent methyltransf. . . IEA.
 DR InterPro; IPR001609; Methyltransf. . .
 DR DR InterPro; IPR000051; SAM_bind.
 DR Hypothetical protein.
 KW SEQUENCE 475 AA; 55147 MW; 0B308C89651526D9 CRC64;

Query Match 16.4%; Score 413; DB 5; Length 475;
 Best Local Similarity 28.6%; Pred. No. 7.3e-24;
 Matches 128; Conservative 70; Mismatches 202; Indels 48; Gaps 14;
 QY 9 AMWDSQASDLKEERPEILSMPLPCKGKLELGGAGIGRETGLAKAGAGVIALDFIBS 68
 DB 24 SMWLNHSAELESSESSADILASLPLLNKQVVDIGAGIGRTTGLAETARWVLSDFIDS 98
 QY 69 AIKQNEVINGHYKQVPMKADVTSPFLSPHSLDVFPSNMLMYLSDEVENLVERMLX 128
 DB 84 FIKKQERNNAHLNINQVGDVAG--LKMSNSVDLVFTNMLMYLSDEETVEFIFNCR 156
 QY 129 WLKPGGYIFFRESCFQHS--GDHKKRS-----NPHYREPRFYTKAFKCHLQDGSNS 180
 DB 157 WLRSHGIVHLRESCEPSTGRSKAKSMHDTANANPHYRFSLSYLINLRAIRYRDVDNKL 216
 QY 181 YELSLSCCKIGAYVRNKNQNOISWLMQVSKDKGQFQFLDTSOYKNSILRYERVF 240
 DB 217 WRFNVQWSCSVPTTYIKRSNNRWQVHWAELKPAEDGAKGTSENLVELIKNTWQNEQAW 276
 QY 241 GPGYVSTGGYETTKFVSMMLDKPGQ-----KVLVGGCIGGGDFYMAETDVEVVG 292
 DB 277 DAKLDEKYYVTDKVFSALTSLSNSSTFLYTPRTVSPYCHINA--HTLAETFNANVMN 334
 QY 293 FDLNWMISFALERSIGLK-CAVEF-----EVAD-CTKINYPDNPSPDVIYSDTLIHQDK 346
 DB 335 TEIIPYRTSLTKSNLKDQVRFGNQSLTDSVTYMQKDALFDVVFATEFLSTVDDE 394
 QY 347 PALFRSPFYKWKPGCKVLISDYCKKAGPPSPFAAYIKORGVDLHDVKEYGOMLKDAGFV 406
 DB 395 --TIRQLPNVMSDGAKEPITL-----PVDEVNEAEMKOR-----IQBLGTYLK--SPT 438
 QY 407 DVLAE--DRTEQPIRLVLRKELETVEKEK 432
 DB 439 DVTDOCTEAOQRYE----KDHEQLRDEK 462

RESULT 12
 Q95PW7 ID Q95PW7 PRELIMINARY; PRT; 484 AA.
 AC Q95PW7;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN ZK622.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCB1_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Leimbach D.;
 RT "The sequence of C. elegans cosmid ZK622.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U39998; AAC008081.1; -
 DR WormPep; ZK622.3b; CE29162.

QY 100 HSLDVFESNMLVLSDEVENLVERMLKMLKPGGVIFPESCFHQSGDHKRS----- 153
DB 4 NSVDLVFTNMLVLSDEVEIFPNCMEWLSHGVLHRESCEPSTGRSKAKSMHDTA 63
QY 154 --NTHVREPRFTYKAFKCHLQDQSGNBYELSLSCKICGAYVKNKQNOISWLAQKV 211
DB 64 NANPHYRPSLSYINLRAIRYDNDVKLWRNFVQWCSVPYIIRKSNWQVHLAEKV 123
QY 212 DSKDDKGFQRFLODTSQYKCNLSILRYRVPFGVSTGGVETTKFPVSMMLDLKPGQ----- 266
DB 124 PAEDCAKGTSPNELVELIKNTWQBOEAWDAKLDEKYVMTDKVPSSALTSLSNSTPFL 183
QY 267 ---KVLVGGCIGGDPYMAETEDVEVGFDSVNMISFALERSIGLK-CAVEP-----EV 318
DB 184 YPRVPSYCHINA--HTLAETFNANWTEIIPYRTSLTKSNLKDQVRFGWQSL 241
QY 319 AD-CTKINYPNSDVIYSRDTLHIQDKPALFRSFKVWLKPGGVLSIDYCKAGPSP 377
DB 242 TDSVTYQCKDALPQVFAVEPLSTVDE--TIRQLPNVMSGAKPITL-----PVDE 293
QY 378 EFAAYIKQGYDLHDVKEYGQMLKADGFDVVLAE--DRTEQPIRVLRKELETVEK 432
DB 294 VNEAEWKQ-----IQBLGYTLK--SFTDVTQCCIEAQGVF-----KHQELRDEK 338

RESULT 15
Q83X76 PRELIMINARY; PRT; 270 AA.
AC Q83X76;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative NBP-hexose 3-O-methyltransferase.
OS Streptomyces rochei (Streptomyces parvullus).
OG Plasmid pSLA2-L.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RA Mochizuki S., Hiratsu K., Suwa M., Ishii T., Sugino F., Yamada K.,
RA Kinashi H.;
RT "The large linear plasmid pSLA2-L of Streptomyces rochei has an
RT unusually condensed gene organization for secondary metabolism.";
RL Mol. Microbiol. 0:0-0(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=20408175; PubMed=10954087;
RA Hiratsu K., Mochizuki S., Kinashi H.;
RT "Cloning and analysis of the replication origin and the telomeres of
RT the large linear plasmid pSLA2-L in Streptomyces rochei.";
RL Mol. Gen. Genet. 263:1015-1021(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=20231737; PubMed=10767533;
RA Suwa M., Sugino H., Sasaoka A., Mori E., Fujii S., Shinkawa H.,
RA Nimi O., Kinashi H.;
RT "Identification of two polyketide synthase gene clusters on the linear
RT plasmid pSLA2-L in Streptomyces rochei.";
RL Gene 246:123-131(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=7434AN4;
RX MEDLINE=99053144; PubMed=9836424;
RA Kinashi H., Fujii S., Hatani A., Kurokawa T., Shinkawa H.;
RT "Physical mapping of the linear plasmid pSLA2-L and localization of
RT the eryA and actI homologs.";
RL Biosci. Biotechnol. Biochem. 62:1892-1897(1998).
DR EMBL; AB088224; BAC76486.1;
GO; GO:0046821; C:extrachromosomal DNA; IEA.

DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransferase; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR004033; UbiE/COQ5_Metrf.
DR Pfam; PF01209; Ubie_methyltran; 1.
KW Methyltransferase; Plasmid; Transferase.
SQ SEQUENCE 270 AA; 29654 MW; F602ECAB8FA82055 CRC64;
Query Match 9.1%; Score 229.5; DB 2; Length 270;
Best Local Similarity 29.4%; Pred. No. 86-10; 83; Indels 23; Gaps 6;
Matches 62; Conservative 43; Mismatches 43;
QY 251 ETTKEFVSMDLKPGQKQVLDVGGIGGGDYMAETFDVEVGFDSVNMISFALERS--I 308
DB 55 QLTQQLARLAPAPGORILDVGGVGPAPFLARTADVWGVSVISAYQVGRAGERARDF 114
QY 309 GLKCAVEFEVADCTKINYPNSDVIYSRDTLHIQDKPALFRSFKVWLKPGGVLSIDY 368
DB 115 GLADRVSPRHADAELPFPDASFDGMAFESLIHMPDKEKVLREIKRVLRFGATLVADM 174
QY 369 CKKAGPSPPEAPYIKQGYDLHDVKEYGQMLKADG-----FVDVLAEDRTQPIRVLRK 423
DB 175 FSQ-----PDTELT-----QDIITPEMADYRAVIESAGLVVRREFTDITATLAPPAYR----- 224
QY 424 ELETVEKE-----KDVFIISDFSEEDYNDIVGG 450
DB 225 --ESVAADLLARKDDVIAMTGPPEFARWVDG 253

Search completed: July 26, 2004, 13:15:27
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2004, 02:21:49 ; Search time 630 Seconds
(without alignments)
3681.246 Million cell updates/sec

Title: US-10-031-331b-40

Perfect score: 2515

Sequence: 1 HTVDLITRAMMLDSQASDLD.....KLRTAKGEQWGLFVAKK 473

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+pin.model -DBV=xlh
-Q=CGN2_1/USPTO_Spool/US10031331/runat_26072004_120354_15090/app_query.fasta_1.647
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blorun62
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10031331 -CGN 1 1 221 -runat_26072004_120354_15090
-NCPU=6 -ICPU=3 -NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
SUMMARIES					

1	2011	80.0	1884	13	US-10-425-114-14370	Sequence 14370, A
2	1394	78.9	1890	13	US-10-425-114-14613	Sequence 14613, A
3	1932.5	76.8	1911	17	US-10-437-963-53228	Sequence 53228, A
4	1829.5	72.7	1962	15	US-10-149-759-81	Sequence 81, Appl
5	1827	72.6	1897	13	US-10-425-114-24087	Sequence 24087, A
6	1734	68.9	2543	13	US-10-424-599-29759	Sequence 29759, A
7	1321	52.5	1201	13	US-10-425-114-36489	Sequence 36489, A
8	1088	43.3	1267	13	US-10-425-114-2412	Sequence 2412, Ap
9	937	37.3	825	13	US-10-425-114-25382	Sequence 25382, A
10	839	33.4	599	17	US-10-021-323-14225	Sequence 14225, A
11	839	33.4	906	13	US-10-424-599-114653	Sequence 114653, A
12	837	31.7	601	17	US-10-021-323-8285	Sequence 8285, Ap
13	798.5	31.7	586	17	US-10-021-323-13961	Sequence 13961, A
14	769	30.6	694	17	US-10-021-323-17175	Sequence 17175, A
15	768	30.5	525	17	US-10-021-323-14357	Sequence 14357, A
16	719	28.6	487	10	US-09-770-961-776	Sequence 776, App
17	690.5	27.5	1087	17	US-10-437-963-7552	Sequence 7552, Ap
18	687	27.3	500	17	US-10-021-323-17176	Sequence 17176, A
19	670.5	26.7	557	13	US-10-424-599-97831	Sequence 97831, A
20	658	26.2	619	15	US-10-149-759-49	Sequence 49, Appl
21	649	25.8	467	12	US-09-732-627A-1204	Sequence 3204, Ap
22	498	19.8	1416	17	US-10-602-268-18	Sequence 18, Appl
23	498	19.8	1534	17	US-10-602-268-6	Sequence 6, Appl
24	484	19.2	434	10	US-09-770-961-584	Sequence 584, App
25	480	19.1	1311	17	US-10-602-268-17	Sequence 17, Appl
26	480	19.1	1533	17	US-10-602-268-5	Sequence 5, Appl
27	428	17.0	272	9	US-09-294-093B-1679	Sequence 1679, Ap
28	427	17.0	526	13	US-10-424-599-81253	Sequence 81253, A
29	422	16.8	539	17	US-10-021-323-17177	Sequence 17177, A
30	421	16.7	685	17	US-10-437-963-7551	Sequence 7551, Ap
31	420	16.7	290	9	US-09-294-093B-5022	Sequence 5022, Ap
32	405	16.1	591	17	US-10-021-323-3964	Sequence 3964, Ap
33	404	16.1	1380	17	US-10-602-268-13	Sequence 13, Appl
34	402	16.0	1786	17	US-10-602-268-1	Sequence 1, Appl
35	399	15.9	477	17	US-10-021-323-6613	Sequence 6613, Ap
36	399	15.9	1380	17	US-10-602-268-14	Sequence 14, Appl
37	399	15.9	1669	17	US-10-602-268-2	Sequence 2, Appl
38	396	15.7	1371	17	US-10-602-268-15	Sequence 15, Appl
39	396	15.7	1472	17	US-10-602-268-3	Sequence 3, Appl
40	390	15.5	293	9	US-09-294-093B-160	Sequence 160, App
41	385	15.3	257	9	US-09-294-093B-2860	Sequence 2860, Ap
42	383	15.2	257	9	US-09-923-876-419	Sequence 419, App
43	383	15.2	257	11	US-09-923-876-419	Sequence 419, App
44	375	14.9	530	17	US-10-437-963-53196	Sequence 53196, A
45	353.5	14.1	1407	17	US-10-602-268-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-10-425-114-14370
; Sequence 14370, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14370
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB22-074-F3_FLI


```

QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
Db 1388 GAGTACATCAACAGACAGAGGATATGATCTCCATGACGTTTATGCACGATGCTA 1447
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspAspThrGluGlnPheLeuArgVal 420
Db 1448 AAGAGCGCTGGCTTCACTGATGTGATCCGACGAGACCGTACTGATCAGTTTATGCAAGTC 1507
QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheLeuSerAspPheSer 440
Db 1508 CTGAAAGTGATTTAGACAGGGTGGAGAAAGAAAGGAAAATTCATCTCCGACTTCTCC 1567
QY 441 GluGluAspTyrIleAspIleValGlyGlyTyrAsnAspLysLeuArgAspThrAlaLys 460
Db 1568 AAGAGAGATTACGATGACATCTGTGAGGATGGAAGTCMAAGCTGGAGAGGTGTGATCG 1627
QY 461 GlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472
Db 1628 GATGAGCAGAAATGGGACTTTTTCATCGCCACAG 1663

RESULT 3
US-10-437-963-53228
; Sequence 53228, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbausk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 53228
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5544C.1
US-10-437-963-53228

Alignment Scores:
Pairing No.: 4,08e-232 Length: 1911
Score: 1932.50 Matches: 359
Percent Similarity: 81.71% Conservative: 52
Best Local Similarity: 71.37% Mismatches: 61
Query Match: 76.84% Indels: 31
DBs: 17 Gaps: 2

US-10-031-331B-40 (1-473) x US-10-437-963-53228 (1-1911)

QY 1 HisThrValAspLeuThrIleGluAlaMetLeuAspSerGlnAlaSerAspLeuAsp 20
Db 82 CACTCCAGGACCTCACCCTGAGGCCATGATGCTCGACTCCGGCGCCGCTCTCGAC 141
QY 21 LysGluGluArgProGlu----- 26
Db 142 AAGGAGAGAGCCCCGAGATCTGGGCGGTGCATGCACTTTTGTGTACTTTTCCCAACC 201
QY 27 -----IleLeuSerMetLeuProProLeuGluGlyLysCysLeu 39
Db 202 AAATACAGTCTCTACGTGGTAAATATGTCTTTACTTCTCTCTACGAGGAGNAATCAGTA 261
QY 40 LeuGluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGln 59
Db 262 CTGGAACTTGTGCTGGATAGTGCCTTCTACTGGAGAACTAGTGAACAACAGCTGGGCAT 321

```

1402 GCTGAAGACCGCACCACATCTCTCGATGTTCTAGAGGAGGAGCTGCTAAGTTGAA 1461
430 LysGluLysAspValPheIleSerAspPheSerGluLysAspValPheValGly 449
1462 AAGACACAAACAGGATGCTCTCTGATTCAGCCAGGAGGACTAGCAGGCCATTTGTAAT 1521
450 GlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGlnArgTyrGlyLeuPheVal 469
1522 GGATGGAGGCAAACTTCAAGAGGATCTGCTGTGAGCAGAGGTGGGGGCTGTTTCATC 1581
470 AlaLysLys 472
1582 GCGACCAAG 1590
RESULT 4
US-10-149-759-81
Sequence 81, Application US/10149759
Publication No. US20030157592A1
GENERAL INFORMATION:
APPLICANT: Lerchl, Jens
APPLICANT: Renz, Andreas
APPLICANT: Ehrhardt, Thomas
APPLICANT: Reindl, Andreas
APPLICANT: Cirpus, Petra
TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
TITLE OF INVENTION: involved in the synthesis of tocopherols and
TITLE OF INVENTION: carotenoids.
FILE REFERENCE: BASF/NAE 1333/99 PCT/US
CURRENT APPLICATION NUMBER: US/10149,759
CURRENT FILING DATE: 2002-10-17
PRIOR APPLICATION NUMBER: PCT/EP/00/12698
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 82
SOFTWARE: WordPerfect version 6.1
SEQ ID NO 81
LENGTH: 1962
TYPE: DNA
ORGANISM: Physcomitrella patens
FEATURE:
NAME/KEY: CDS
LOCATION: (367)..(1842)
OTHER INFORMATION: 78 ppprot1_092_e12-260rev
US-10-149-759-81
Alignment Scores:
Pred. No.: 3.9e-219 Length: 1962
Score: 1829.50 Matches: 332
Percent Similarity: 83.76% Conservative: 65
Best Local Similarity: 70.04% Mismatches: 74
Query Match: 72.74% Indels: 3
Gaps: 2
DB:
US-10-031-331b-40 (1-473) x US-10-149-759-81 (1-1962)
Qy 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
Db 415 CATTCGTGAGGAGCCTAGGCAATGATGCTTGTGAGGAGGCTCCAAACTCGAT 474
Qy 21 LysGluLysArgProGluLysLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40
Db 475 AAGAGAGACGACCCGAGATTTGTGCTGTGGCCATATGAAACACAGGATGTCATG 534
Qy 41 GluLeuGlyValAlaGlyIleGlyValGlyPheThrGlyGluLeuAlaGlyLysValGlnVal 60
Db 535 GAGCTCGAGGAGGAGCATCGTTCGTGTTACTGTGTGCTTGTGAGGATGCGAGTTCATG 594
Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
Db 595 CTTCGATGATTTTCATGAGGATCTCATCAAGAGACGAGGATGTGAACGGTCACTAC 654
Qy 81 LysAsnValLysPheMetCysAlaLeuValThrSerProThrLeuSerPheProHis 100

60 ValIleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHis 79
322 GTTCTGTCATGATGATTTCTAAGAGTGTGATTAAGAGATGAACATATAAGGCTCAC 381
80 TyrLysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPhePro 99
382 CACAGAATGATCATCTTTATGTTGTCGATGTCATCTCCAGACCTGATGATGAGGAT 441
100 HisSerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluVal 119
442 AACTCCATGATCTGATATTTTCAAACTGCTTACTGATGATCTTTTCAGACGAGGAT 501
120 GluAsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArg 139
502 GAGAGCTAGTAAGAGATGTAAGATGCTTAAGGTTGGCGCTATATCTCTTTAGG 561
140 GluSerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArg 159
562 GAATCTTTGTTCCATGCTGAGATTCANAAAGAGATGAATCTTACATATACCGG 621
160 GluProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsn 179
622 GAGCAAGGTTTTACATCAAGGTGTTTAAAGAGTGTCAAGCTCTTGATCAAGATGGGAT 681
180 SerTyrGluLeuSerLeuLeuSerCysLysCysIleGlyValTyrValArgAsnLysLys 199
682 TCCCTTGAATCTCTGCTACTTCTGAGTGTGTTGAGCTTACGTGAAAGCAAGAAA 741
200 AsnGlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLysGlyPhe 219
742 AATCAAAACCATGATATGTTGGCTATGTCGCAAAAGTTGATTCACAGAGATCGGGGTT 801
220 GlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgVal 239
802 CAAGATTTTGGCAATGTGCTATCAAGCCAGTGCATATATACCTATCAAGCATC 861
240 PheGlyProGlyTyrValSerThrGlyTyr----- 250
862 TTTCGAGAGGCTTTGTGAGCAGCTGGTGGATTTGTATGCTATCTTTTATCTTCGTCA 921
251 ---GluThrThrLysGluPheValSerMetLeuAspLeuLysProGlyLysValLeu 269
922 CCAGAACTACAAAGATTTGTGACAGGCTGATCTCAAACTCGCCAGAACGTTCTT 981
270 AspValGlyCysGlyIleGlyGlyValAspPheTyrMetAlaGluThrPheAspValGlu 289
982 GATGTTGGATGTTGGATTTGGGGGGTGAATTTTATATGCTGCACAGTATGATGTTCA 1041
290 ValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGly 309
1042 GTTGTGTTGATGATCTTTCGATAAACATGTTTCTTTTGCACCTGAGCGTCTATTTGG 1101
310 LeuLysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsn 329
1102 CGTAAGTGTCTGATTTGAGTTGAGTGTGCTGATTCACCAAAAGACATACCCAGACAAC 1161
330 SerPheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeu 349
1162 ACCTTTGACGATCTACAGTGTGATGATCTATCTTACATACAGATTAACCTCACTA 1221
350 PheArgSerPheTyrLysTrpLeuLysProGlyLysValLeuIleSerAspTyrCys 369
1222 TTTAAAGTTTCTTCAGTGGCTCAAACTCGGGGTAGGTCCTTAATAGTGAATCTGCT 1281
370 LysLysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAsp 389
1282 AAGTGCCTGGGAAACCTTCAGAGAGTTCGAGCTTACATTAAGCAAGGGTTATGAC 1341
390 LeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeu 409
1342 CTTCCAGCGTCAAGGCTTACGACAGATGCTTGAGATGCTGTTCATGATGATGAT 1401
410 AlaGluAspArgThrGluGlnPheIleArgValLeuArgLysGluLeuGluThrValGlu 429

```

Db      655 AACACATCGATTTCAATGTGGGATGTGACCTCTCCAGACCTGATATATTCAGCAGGT 714
QY      101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluValGlu 120
Db      715 TCTGGGATCTCGTGTCTTCAATTTGGCTTCTCATGTACTTGTCTGACGAGAGGTAAA 774
QY      121 AsnLeuValGluArgMetLeuTyrTrpLeuPheProGlyTyrIlePhePheArgGlu 140
Db      775 GCTTTAGCATCAGCGGTATGATGCTGCTGAGCTGGAGATACATTTTCTTCAGAGAA 834
QY      141 SerCysPheHisGlnSerGlyPheHisArgIleGlySerAsnProThrHisTyrArgGlu 160
Db      835 TCTCTCTTCCACCTAGTACGAGATCAACAGCGAAGAAACAATCTACTCACTACCGTCAA 894
QY      161 ProArgPheTyrTrpLysAlaPheLysGluCysHisLeuGlnAspGlySerGlySerAsn 180
Db      895 CCAACAGAGTACAGAACATCTTCCAGCAGCGCTACATCGAGAG---GATGGCTCTAT 951
QY      181 TyrGluLeuSerLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200
Db      952 TTCAGGTTGAATGGTCGGATGCCAATGTGCGCACATACGTGCGAATAAGAAAT 1011
QY      201 GlnAsnGlnIleSerTrpLeuTrpGlnLysVal-----AspSerLysAspLysGly 218
Db      1012 CAAACACAGGTGTGTGTGTATGAGGAAAGTTTCAGTCGCGATCGACCTGACGAGCGTGT 1071
QY      219 PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg 238
Db      1072 TTCAGAGATTTTGGACACCCACAGTACAGTCACTGGAATCTCGGTACGAGCGT 1131
QY      239 ValPheGlyProGlyTyrValSerThrGlyTyrGluThrThrLysGluPheValSer 258
Db      1132 ATTTTGGAGAGAGTTTGTAGCAGGCTGGATCGAATCGAAGAGCTTTGTAGT 1191
QY      259 MetLeuAspLeuLysProGlyGlnLysValLeuAspValCysGlyIleGlyGly 278
Db      1192 ATGCTGACCTTGAAGCCAGCAGCAGCTGCTTGTGACGTTGGATGGGATCGAGGTGT 1251
QY      279 AspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsn 298
Db      1252 GATTTCTACATGCGCGAGAGATATGATGTGAAGTTGTGGCATCGACTGTCTTAAT 1311
QY      299 MetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluVal 318
Db      1312 ATGATTTCTGTCTTGAACATCGATCGGACAGAAATGTGCACTGAGTGTGAAGT 1371
QY      319 AlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAsp 338
Db      1372 GGGGATTCACCAAGATTAATACCTCAGCGCATCTTTTGTATGATCATCTACAGTCGTGAT 1431
QY      339 ThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLys 358
Db      1432 ACCATTCACATTCAGATTAACCTCGGCTTTTTCACACGTTTATTAATGTTGAAG 1491
QY      359 ProGlyCysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGlu 378
Db      1492 CCTGGAGTGGGTGCTTATCAGTCACTACTGATAGCTCCACAACTCCGTCGCGGAG 1551
QY      379 PheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGln 398
Db      1552 TTGCTGCATACATTCCAGCAGGGGTATGATCTCCATACCGTTTATTAATGTTGAAG 1611
QY      399 MetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIle 418
Db      1612 ATGCTGAAGATCCCGTTTGTGAAGTGTGCGAGAGGACCGCAGCATGATTCATT 1671
QY      419 ArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAsp 438
Db      1672 GAAGTGTTCACAGAGGAGTACCCACCTAGACGAGCTGTGACCGATTCATCACAGAT 1731
QY      439 PheSerGlnIleAspTyrAsnAspIleValGlyTyrAsnAspLysLeuArgThr 458
Db      1732 TTCTCCAGAGAGATTATAACTACATTGTGAGCGGATGAGAGTAACTGAGCGCTGT 1791

```

```

QY      459 AlaLysGlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
Db      1792 TCGAATGACGACAGAGTGGGACTCTTCATAGCCTACAG 1833

```

RESULT 5

```

US-10-425-114-24087
; Sequence 24087, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jiqdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 24087
; LENGTH: 1897
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3607-019-D6_FLI
US-10-425-114-24087

```

```

Alignment Scores:
Pred. No.: 7,62e-219 Length: 1897
Score: 1827.00 Matches: 341
Percent Similarity: 83.47% Conservative: 53
Best Local Similarity: 72.25% Mismatches: 78
Query Match: 72.64% Indels: 1
DB: 13 Gaps: 0
US-10-031-331B-40 (1-473) x US-10-425-114-24087 (1-1897)

```

```

QY      1 HisthrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
Db      240 CACTCCGGGAGCTCAACCTCGAGGCCATTATGCTGACTCCCGCGCCGCGAATCGAC 299
QY      21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40
Db      300 AAGGAGAGCGCCCGAGGTTCTGTCTTACTTCTTCATATGAGGGAATCTACTG 359
QY      41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
Db      360 GAGCTGGAGCTGGAATAGCGCGCTTACTGTGAACTGGCTTAACATCTGGGCGATGTT 419
QY      61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
Db      420 TTTCAGTGTGATTCGTTGAAAGTGTGATTAAAGATGGAAGTATAAATGATCACTAT 479
QY      81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
Db      480 GGCAACACATCTTTATGTCTGATGTATACATCCCGGACCTGATGATGAGCAAC 539
QY      101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120
Db      540 TCAATTGATCTGATTTTCAAACTGGTGTGCTGATGATCTTTTCAGATGAGGAGATTGAC 599
QY      121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheArgGlu 140
Db      600 AAGTTGTGTAAGAAGATGTTAAATGGTTGAAGTGGTGTATTCTTCTTTAGGGA 659
QY      141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
Db      660 TCTTGTCTTCATCANTCCGAGATACAGAAAGGAATTTTATCCACACATCTACGAGAA 719
QY      161 ProArgPheTyrThrIysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180

```


1313 GATGATATGCGCCAGGCTTTGTGAGCACTGTGTGAAACGACAAAGAGATTGTGG 1372
 258 erMetLeuaspLeuysProGlyGlnlyValLeuaspValGlyCysGly11leGlyGlyG 278
 1373 CAAAGCTGGAGCTAAAGCTGTGAGAAAGTTCTGGATGTGTGGTGTGTGGGAG 1432
 278 lyAspPheTyrMetAlaGluThrPheaspValGluValGlyPheaspLeuSerVala 298
 1433 GTGATTTCTACATGCTGAGAATTTTGATGTTGAGGTTATTTGGCATTGACCTCTCATAA 1492
 298 snMetIleSerPheAlaLeuGluArgSerileGlyLeuysCysAlaValGluPheGluV 318
 1493 ACATGATTTCTGCTATTGAGCTGCTATTGAGCTGAACTAGCGCTGTGTGAATTGAT 1552
 318 alAlaaspCysThrlyIleAsnTyrProaspHsnSerPheaspValIleTyrSerArgA 338
 1553 GTGCGGATTTGATAGAAACATATCTCTGAGATACATTTGATGATATATACACCTCGG 1612
 338 spThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrlystToLeuL 358
 1613 ACACCATGCTGCAGTCAAGATTAACACATATTTAGATCATTTTACAAAGTGTGGA 1672
 358 yspProGlyGlyIleValLeuIleSeraspTyrCyslyLysAlaGlyProProSerProG 378
 1673 ACCTGGAGGAAATTTCTATCATCAGATTTACTGCAAAAGTGTGGAAGTCCATCTTAG 1732
 378 luPheAlaIleTyrIleLysGlnArgGlyTyrAspLeuHisaspVallysgLutvrglyG 398
 1733 AATTTGCTGAGTACATAAAAGAGGAGATTTATCTCCATGACATTAAGCGGTATAGC 1792
 398 InMetLeuLysAspAlaGlyPheValaspValLeuAlaGluaspArgThrGluGlnPheI 418
 1793 AGATGCTTGAGATGCGGATTTGATGATGTCATGTCGAGAGATCGAATGATGATGTTG 1852
 418 leArgValLeuArgLysGluLeuGluThrValGluLysGlyLysaspValPheIleSerA 438
 1853 TGAACACGCTACAGCAGGAGTAAATGCCCTTGAGAACAGAGGAGCGATTTATTGCTG 1912
 438 spPheSerGluGluaspTyrAsnaspIleValGlyGlyTyrAsnaspLysLeuArgArg 458
 1913 ACTTCAGCGAGAGAGATACATGAGATTTGCAAGATGGAAGCGAAGCAGACCGAGG 1972
 458 hrAlaLysGlyGluGluArgTyrGlyLeuPheValAlaLysLys 472
 1973 GTGCATCTAGAGCAGATGCGGGCTTGTTCATTGCCAAGAA 2016

RESULT 7

US-10-425-114-36489
 / Sequence 36489, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jingdong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Screen, Steven E
 / APPLICANT: Tabaska, Jack E
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / CURRENT APPLICATION NUMBER: US/10/425,114
 / FILE REFERENCE: 38-21(53313)B
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 36489
 / LENGTH: 1201
 / TYPE: DNA
 / ORGANISM: Zea mays subsp. mexicana
 / FEATURE:
 / OTHER INFORMATION: Clone ID: UC-ZMOT0505INTE108B03_PL1
 US-10-425-114-36489

Alignment Scores:

Pred. No.: 1.74e-155 Length: 1201

Score: 1321.00 Matches: 241
 Percent Similarity: 87.26% Conservative: 33
 Best Local Similarity: 76.75% Mismatches: 40
 Query Match: 52.52% Indels: 0
 DB: 13 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-425-114-36489 (1-1201)

QY 159 ArgGluProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnaspGlySerGly 178
 Db 1 CGAGAACCAAGGTTTATACCAAGGTATTAAAGAGGCGCATTCATTTCATCAGATGGA 60
 QY 179 AnSerTyrGluLeuSerLeuSerCysLysCysLysCysLysCysLysCysLysCys 198
 Db 61 GGTTCGTTTGAACCTTCTAGTGACCTGTAAATGATTTGGGCTTATGTGTCAAAACAG 120
 QY 199 LysAsnGlnAnGlnIleSerTyrLeuTyrGlnLysValaspSerLysaspLysGly 218
 Db 121 AAGAAATCAAAACAGATATGCTGTTATGGGAAAGGTAAATCAACAGACAGAGAT 180
 QY 219 PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg 238
 Db 181 TTTCAAGATTTCTGGACACGTCATACAAACAGTGGGATATTACGTTATGAGGT 240
 QY 239 ValPheGlyProGlyTyrValSerThrGlyTyrGluThrThrlysgLupheValSer 258
 Db 241 GTCTTTGTGAGGTTTGTGAGCCTGTGGGATCGAGACTACAAAGGAAATTTGTGGGC 300
 QY 259 MetLeuaspLeuysProGlyGlnLysValLeuaspValGlyCysGlyIleGlyGly 278
 Db 301 ATGCTCGATCTTAAACCGGCGCAGAAAGTACTTGATGTCGATGTGGAATTTGAGCGGC 360
 QY 279 AspPheTyrMetAlaGluThrPheaspValGluValGlyPheaspLeuSerValasn 298
 Db 361 GACTTTACATGCTGCAAACTATGATGTCATGTTCTTGGTATTGATCTTTGGTGAAC 420
 QY 299 MetIleSerPheAlaLeuGluArgSerIleGlyLysCysAlaValGluPheGluVal 318
 Db 421 ATGCTTCATTTGCAATTTGAACGTCGCAATTTGAGCGCAAGTCTCTGTTGAATTCGAGTT 480
 QY 319 AlaaspCysThrlyIleAsnTyrProaspAnSerPheaspValIleTyrSerArgasp 338
 Db 481 GCTGATTCACCAAGGATTAACCCAGAAATAGTTTTCAGCTCATCTACGCCGTGAC 540
 QY 339 ThrIleLeuHisIleGlnaspLysProAlaLeuPheArgSerPheTyrLysTyrLeuLys 358
 Db 541 ACCATCTTCACATACAGACAAAGCTGCTCTGTTTCAAGAGCTTCTTCAAAATGCTAAAG 600
 QY 359 ProGlyLysValLeuLysSeraspTyrCysLysLysAlaGlyProProSerProGlu 378
 Db 601 CCGCGCGCAAGTCTTAATCAGCGACTACTGTGAAGATCTCTGGAACCATCAGAGAA 660
 QY 379 PheAlaIleTyrIleLysGlnArgGlyTyrAspLeuHisaspVallysgLutvrglyGln 398
 Db 661 TTTGCTGCTACATTAAAGCAGAGAGCTATGACCTTCAGAGCTGAAGGCTTATGGACAG 720
 QY 399 MetLeuysaspAlaGlyPheValaspValLeuAlaGluaspArgThrGluGlnPheIle 418
 Db 721 ATGCTGAAGGATGCTGTTTTCATATATGTCATCGCGAAGATGCGACTGACGAGTCTCTG 780
 QY 419 ArgValLeuArgLysGluLeuGluThrValGluLysGlyLysaspValPheIleSerasp 438
 Db 781 AATGTTCTCAGAGGAGATAGGTGAGATTTGAAAGAAACAAAGACGCTTCTCTGGCAGAC 840
 QY 439 PheSerGluGluaspTyrAnaspIleValGlyGlyTyrAnaspLysLeuArgArgThr 458
 Db 841 TTCACCGAGAGGACTATGACGACATTTGTGAATGCTCGAACCGGAGCTGAAACGAGC 900
 QY 459 AlalysGlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472
 Db 901 TCTGCCGCGAGCAGAGGTGGGGTGTGTTTCATTGCCACCAAG 942

RESULT 8

US-10-425-114-2412
 ; Sequence 2412, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 2412
 ; LENGTH: 1267
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700211781_FLI
 ; US-10-425-114-2412

Alignment Scores:
 Pred. No.: 3 98e-126 Length: 1267
 Score: 1088.00 Matches: 210
 Percent Similarity: 76.54% Conservative: 38
 Best Local Similarity: 64.81% Mismatches: 68
 Query Match: 43.28% Indels: 8
 DB: 13 Gaps: 3

US-10-031-331b-40 (1-473) x US-10-425-114-2412 (1-1267)

QY 153 SerAspProThrHis-----TyrArgGluProArgPheTyrThrIysAlaPhe 168
 DB 34 TCAAGCCAACTATCTGGTTTGGATCTTCTTGGATGTTTACATAAGCTCCAAATGCTT 93
 QY 169 LysGluCysHisLeuGlnAspGlySerGlyAsnSerTyrGluLeuSerLeuSerCys 188
 DB 94 ---GATGTAATCAAGAAAGTTTGAAGAAAGTCCATC-----CTGATTAAAGT 141
 QY 189 LysCysIleGlyAlaTyrValArgAsnIysAsnGlnAsnGlnIleSerTyrLeuTyr 208
 DB 142 TTGGCATCTTTAAATACCTTGGTATATAAACTTGGTTCTCGATAGTGTGGCTATGG 201
 QY 209 GlnIysValAspSerIysAspAspIysGlyPheGlnArgPheLeuAspThrSerGlnTyr 228
 DB 202 AAAAAGTAATCATCATCAGACAGATGGGGATTTTCAAAGTTTTTGGCAATGTGCAATAC 261
 QY 229 LysCysAsnSerIleLeuArgTyrGluArgValPheGlyProGlyTyrValSerThrGly 248
 DB 262 AAAGCCACTGGAATACCTACGCTATGAACGTATCTTTGGAGATGGCTACGTGAGTACTGT 321
 QY 249 GlyTyrGluThrThrIysGluPheValSerMetLeuAspLeuIysProGlyGlnIysVal 268
 DB 322 GGAGCTGAGACTACAAAGAAATTTGTGGGAAACTGATCTTAAGCTCTGGGCAAGAGTG 381
 QY 269 LeuAspValGlyCysGlyIleGlyAspPheTyrMetAlaGluThrPheAspVal 288
 DB 382 CTTGATGCTTGGATGTGGAAATTTGGGGAGAGTGACTTTTATATGCTGGAAGATATGTGACA 441
 QY 289 GluValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIle 308
 DB 442 CATGTTGTTGGTATTGACCTTCTCATACATGATAAATGTTTGGCTTGGAGCGTTCATTT 501
 QY 309 GlyLeuIysCysAlaValGluPheGluValAlaAspCysThrIysIleAsnTyrProAsp 328
 DB 502 GGGTGAAGTCTTAGTGTGAGTTTGAAGTTTGGATGCTGATCCACCAACAGACATACCCAGAC 561
 QY 329 AsnSerPheAspValIleTyrSerArgAspThrIleLeuHsrIleGlnAspLysProAla 348
 DB 562 CACATGTTGATGTCATCTACATGCTGACACTCTCTTCATATACAGATAAACCTCTCC 621

QY 349 LeuPheArgSerPheTyrIysTyrLeuIysProGlyGlnIysValLeuIleSerAspTyr 368
 DB 622 TTGTTTAAAGAGTTTCTTCAATGGCTGAAACCTGGGGGAAAGGTTCTTAATCAGTGATTAC 681
 QY 369 CysIysIysAlaGlyProSerProGluPheAlaIleTyrIleLysGlnArgGlyTyr 388
 DB 682 TGCAAGAGTCTCTGGAAGAACCATCAGAAAGAGTTTGCACATACATATTAAGCAGAGGGGTAT 741
 QY 389 AspLeuHisAspValIysGlnTyrGlyGlnMetLeuIysAspAlaGlyPheValAspVal 408
 DB 742 GATCTCATGATGTGGAGGCTTATGACAGATGCTCAGATGCTGGTTTCAGTCATGTC 801
 QY 409 LeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgIysGluLeuGluThrVal 428
 DB 802 ATTGCTGAAGACCGAACTGACCACTTCTCAGTGTGTTTACAGAGGAGCTAGACAAATTT 861
 QY 429 GluIysGluIysAspValPheIleSerAspPheSerGluGluAspTyrAsnAspIleVal 448
 DB 862 CAGAAGAACAAAGATGATTCTCTGAGTTTGGCCAGGAGGATTTATGACGATATCGT 921
 QY 449 GlyGlyTyrAsnAspIysLeuArgArgThrAlaIysGlyGluGlnArgTyrGlyLeuPhe 468
 DB 922 AATGATGTGAAGGCAAACTGACAGAGCTCTGCTGTGAGCAGAGGTGGGGGCTGTTTC 981

RESULT 9
 US-10-425-114-25382
 ; Sequence 25382, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 25382
 ; LENGTH: 825
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3829-031-P4_FLI
 ; US-10-425-114-25382

Alignment Scores:
 Pred. No.: 2 01e-107 Length: 825
 Score: 937.00 Matches: 176
 Percent Similarity: 88.07% Conservative: 16
 Best Local Similarity: 80.73% Mismatches: 26
 Query Match: 37.26% Indels: 0
 DB: 13 Gaps: 0

US-10-031-331b-40 (1-473) x US-10-425-114-25382 (1-825)

QY 255 GluPheValSerMetLeuAspLeuIysProGlyGlnIysValLeuAspValGlyCysGly 274
 DB 2 GAAATTTGTGGCAAGTTGGATCTAAAGCTCTGGCAAGGTTCTAGAGTTGGCTGTGGC 61
 QY 275 IleGlyCysIleAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAsp 294
 DB 62 ATAGGTGAGGTGACTTTTATATGTCAAAGAAATTTGGAGTCTATGTTGTGGGCTCGAT 121
 QY 295 LeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuIysCysAlaVal 314

237 GluArgValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPhe 256
1 GAGAGTGTGTTTGGTCAAGGTTTGTGAGCAGAGGAGGACTTGAACAACACCAAGCAATTC 60
257 ValSerMetLeuAspPheGlyProGlyGlnLysValLeuAspValGlyCysGlyIleGly 276
61 GTGGCAAGTTGGGACTTAACCTGGCCGAGAGTACTGGATGTTGGTGTGGTACTGGG 120
277 GlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSer 296
121 CGAGGTGCATTTACATGGCAGAAATTTTGTATTTGAGGTTGTGGCATTTGACCTCTCC 180
297 ValLeuMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPhe 316
181 ATAAACAATATTTCTCTGGCAATTTGAACGTTGATTTGGACTCAATGCTGTGTGAATTT 240
317 GluValAlaAspCysThrIleAsnTyrProAspAsnSerPheAspValIleTyrSer 336
241 GAGTGTGGGATTTGACATAAACAATTTCTGTGATACATTTGATGATATCTATTTCC 300
337 ArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrp 356
301 CGTGACACATTTGTACATCAATCAAAAGATAAGCCATCTATTTCAATCATTTTACAAATGG 360
357 LeuLysProGlyGlyLysValLeuIleSerAspTyrCysAlaLysAlaGlyProProSer 376
361 TTGAGCGTGGAGTACATCTCTGATTTACTGATTTACTGATTTACTGCAAAAGTGAAGTCTATCA 420
377 ProGluPheAlaAlaTyrIleGlnArgGlyTyrAspLeuHisAspValLysGlyTyr 396
421 TTAGGATATGCTGAGTATATAAAGAGGGGGATATATATTATCATGATGAAACATAT 480
397 GlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGln 416
481 TGTGGATGCTGAGATGCTGGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 540
417 PheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIle 436
541 TTATGAAACACTACACAGAGGTTAATGCTTCCACAGCAGAGGATGATGATGATGATGAT 600
437 SerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTrpAsnAspLysLeuArg 456
601 GATGACTTCTCGAGGAGAGACTACAAATGAAATTTACTGAAAGATGGAAGCCAGCAGATG 660
457 ArgThrAlaLysGlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473
661 CGGGGGCGAGTGGTGAACAAATATGAGCTTGTTCATTCGCAAGGAAGAA 711

RESULT 12
US-10-021-323-8285
; Sequence 8285, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Peng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 8285
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-008-Q1-K6-86
US-10-021-323-8285

Alignment Scores: 4.62e-95 Length: 601
Pred. No.: 837.00 Matches: 162
Score: 88.38% Conservative: 13
Percent Similarity: 81.82% Mismatches: 23
Best Local Similarity: 33.28% Indels: 0
Query Match: 17 Gaps:
DB: 1
US-10-031-331B-40 (1-473) x US-10-021-323-8285 (1-601)
QY 254 LysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCys 273
DB 8 AAAAGATTGTGGGGAAGTTAGATCTTAAGCCCTGGCCAAAGTC-CTAGATGTTGGCTGT 66
QY 274 GlyIleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPhe 293
DB 67 GGCATTGTGGAGGTGACATTTATATGGCTGAGCAATTTGATGTTTCATGTTGTGGGCAAC 126
QY 294 AspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeuLysCysAla 313
DB 127 GACCTCTCTGTAAACATGATATCTTCTCTTGAACGAGCTACTGGACTGGAATGCTCA 186
QY 314 ValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspVal 333
DB 187 GTGGAAATTTGAAGTTGCTGATTCACCAAGAGTTTATCCGACACACAGTTTGTGTT 246
QY 334 IleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPhe 353
DB 247 ATCTACACCGGTGACACTATTTCTACATTCATGACAAACCTGCACCTGTTAGATCTTTC 306
QY 354 TyrLysTrpLeuLysProGlyGlyLysValLeuIleSerAspTyrCysAlaLysAlaGly 373
DB 307 TACAATGTTTGAAGCAGGAGGCAATCTCTCATAGTGATTTACTGCAAAAGTTCCAAAG 366
QY 374 ProProSerProGluPheAlaAlaTyrIleGlnArgGlyTyrAspLeuHisAspVal 393
DB 367 ACTCCATCCAGGAGTTTGTCTGAGTATATCAACGACAGAGGCTATGATCTTCATGATGA 426
QY 394 LysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArg 413
DB 427 AAATCATATGACAGATGCTTGAGAGTCTGTTTGTATGATGATTTCTTCACAGGATCGA 486
QY 414 ThrGluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGlyLysAsp 433
DB 487 ACCGATCAGTCTTCAAGTTCTCGGCGTGAATTTGAACCAAGTGGAGAAAGAGAGGAT 546
QY 434 ValPheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyGlyTrp 451
DB 547 GCATTCATCTCTGACTTCTCTAAGGAGAGACTATGATGAAATAGTTGGTGGATGG 600

RESULT 13
US-10-021-323-13961
; Sequence 13961, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Peng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 13961
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure

LOCATION: (1) (586)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3829-022-Q6-K6-H6
US-10-021-323-13961

Alignment Scores:
Pred. No.: 3 11e-90 Length: 586
Score: 798.50 Matches: 150
Percent Similarity: 85.56% Conservative: 26
Best Local Similarity: 80.21% Mismatches: 1
Query Match: 31.75% Indels: 1
DB: 17 Gaps: 1

US-10-031-331B-40 (1-473) x US-10-021-323-13961 (1-586)

QY 106 PheSerAsnTTPLeuMetTyrLeuSerAspGluValGluAsnLeuValGluArg 125
DB 26 TTTCTCACTGGCTGTTGATGATATCTCTCANTAAAGAGTTGAGAA---TTTGGCGAAGG 82
QY 126 MetLeuLysTTPLeuLysProGlyGlyTyrIlePhePheArgGluSerCysPheHisGln 145
DB 83 ATGCTCAAGTGGTTGAAGGTTGGTGACATATTTCTTCAGGAGTCTTGTCTTCATCA 142
QY 146 SerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGluProArgPheTyrThr 165
DB 143 TCTGGTGATTGCAAGAGAAAGACACCACTCACTTCGCGAGCCAGATTTTCACT 202
QY 166 LysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSerTyrGluLeuSerLeu 185
DB 203 AGGTCTTTAAAGATGCCAAGCAACTGATGATTCGGAATTCATTGAACTTTCTCTT 262
QY 186 LeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsnGlnIleSer 205
DB 263 GTGGCTGCAAGTCATGGAGCTTATGTTAAGAACAAAGAAATCAGAAATTCAGTTTC 322
QY 206 TTPLeuTTPGlnLysValAspSerLysAspAspGlyPheGlnArgPheLeuAspThr 225
DB 323 TGGATCTTCCAAAAGTTGTTTCAGATATGATAGGATTCAGGCTCTTGTGATTC 382
QY 226 SerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheGlyProGlyTyrVal 245
DB 383 GTTCAGTCAAACTTAATAGCAATCTCCGTTATAGCGGTCTCTTGGACAGGATATGTG 442
QY 246 SerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeuAspLeuLysProGly 265
DB 443 AGCACAGGAGGAATGAAACAAACAAAGAAATTTGTGGGAAGTTAGATCTTAAGCTGGC 502
QY 266 GlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPheTyrMetAlaGluThr 285
DB 503 CAAGAGTCTAGATGTGGCTGTGGCATTTGGTGGAGGTGACATTTATATGCTGAGGA 562
QY 286 PheAspValGluValGly 292
DB 563 TTTGATGTTTATGTTGGGC 583

RESULT 14

US-10-021-323-17175/c
Sequence 17175, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255,619
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 17175

LENGTH: 694

TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3829-031-Q6-K6-F4-PPOLYA
US-10-021-323-17175

Alignment Scores:
Pred. No.: 2 11e-86 Length: 694
Score: 769.00 Matches: 144
Percent Similarity: 86.89% Conservative: 15
Best Local Similarity: 78.89% Mismatches: 24
Query Match: 20.58% Indels: 0
DB: 17 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-021-323-17175 (1-694)

QY 290 ValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGly 309
DB 692 GTTGTGGCATCGATCTGTCATACATATCAATCTTTGCTCTTGAAGAGCCATGCA 633
QY 310 LeuLysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspLeu 329
DB 632 CTCGATTCTCTGTGGAATTTGAAGTTGCTGATTCGACTACAAAGACTTATCCGATAAC 573
QY 330 SerPheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeu 349
DB 572 AGCTTTGATGTTATCTACAGCCGTCACAGATCTACACATTCACAGACAGCTGCACTA 513
QY 350 PheArgSerPheTyrLysTTPLeuLysProGlyGlyValLeuLysSerAspTyrCys 369
DB 512 TTTAGATATTTCTACAAATGTTGAAGCCAGGGGCAAGTTCTTATAGTGTATTACTGC 453
QY 370 LysLysAlaGlyProProSerProGluPheAlaIleTyrIleLysGlnArgGlyTyrAsp 389
DB 452 AAAGCCGAGAGGCTTTCATGCCGGAATTTGACAGATACATCAAGCAGAGGCTATGAT 393
QY 390 LeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeu 409
DB 392 CTCATGATGTTAAAGCTTATGGCCAGATGCTTAAGATGCTGTTTGTATCAGTCAAT 333
QY 410 AlaGluAspAspThrGluGlnPheIleArgValLeuArgLysGluLeuThrValGlu 429
DB 332 GCAGAGATTCGACAGATGATGATCATACAGTTCTCAGAGGGAATTAGACAAAGTTGAG 273
QY 430 LysGluLysAspValPheIleSerAspPheSerGluAspTyrAsnAspIleValGly 449
DB 272 AAGGACAGGATACCTTTCATCCGACTTTTCTCAGAGAGACTACGACGATATCGTTGG 213
QY 450 GlyTTPAsnAspLysLeuArgAspThrAlaLysGlyGluGlnArgTTPGlyLeuPheVal 469
DB 212 CGATCGAAGGCAGAGCTTAATAGGACCGAGTTTGGGAGCAGAGATGGGCTCTTTTCAAT 153
QY 470 AlaLysLys 472
DB 152 GCCAACAAA 144

RESULT 15

US-10-021-323-14357
Sequence 14357, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14

; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14357
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Cossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-033-Q1-K6-A5
US-10-021-323-14357

Alignment Scores:
Pred. No.: 1.79e-86 Length: 525
Score: 768.00 Matches: 146
Percent Similarity: 89.66% Conservative: 10
Best Local Similarity: 83.91% Mismatches: 18
Query Match: 30.54% Indels: 0
DB: 17 Gaps: 0

US-10-031-331B-40 (1-473) x US-10-021-323-14357 (1-525)

QY	251	GluThrThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAsp	270
DB	4	GAACACACAAAGAAATTGTGGGAGGTAGATCTTARGCCTGGCCAAAGGCTCTAGAT	63
QY	271	ValGlyCysGlyIleGlyGlyLysPheTyrMetAlaGluThrPheAspValGluVal	290
DB	64	GTGGCTGTGGCATTTGTGGAGGTGACATTTATATGCTGAGGAATTTGATGTTTCATGTT	123
QY	291	ValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArgSerIleGlyLeu	310
DB	124	GTGGGCAACGACCTCTCTGTAAACATGATATCTTTGCTCTTGAACGAGCTACTGACTG	183
QY	311	LysCysAlaValGluPheGluValAlaAspCysThrIleAsnTyrProAspAsnSer	330
DB	184	AAATGCTCAGTGGAAATTTGAAGTTGCTGATTCACCAAGAGGTTTATCCGGACACAGT	243
QY	331	PheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPhe	350
DB	244	TTTGATGTATNCTACAGCCGTGACACTATTCTACACATTCATCACAACCTGCACTGTTT	303
QY	351	ArgSerPheTyrLysThrPheLysPheGlyGlyLysValLeuIleSerAspTyrCysLys	370
DB	304	AGATCTTTCTACAAATGTTGARGCCAGAGGCAAACTCCTCATAGTGAATTAAGTGCATA	363
QY	371	LysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu	390
DB	364	AGTTCACAGACTCCATCCAGGGAGTTTGTCTGAGTATATCAAGCAGAGAGGCTATGATCTT	423
QY	391	HisAspValLysGluTyrGlyGluMetLeuLysAspAlaGlyPheValAspValLeuAla	410
DB	424	CATGATGTAAATCATATGACAGATCTTGAGGATGCTGTTGATGTGATTTGATGATTTGCA	483
QY	411	GluAspArgThrGluGlnPheIleArgValLeuArgLysGlu	424
DB	484	GAGGATCGAACCGATCAGTTCTTACAGTTCTGCGGGGTGAA	525

Search completed: August 2, 2004, 05:04:18
Job time : 651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 2, 2004, 00:38:15 ; Search time 102 Seconds

Title: US-10-031-331B-40
Perfect score: 2515
Sequence: 1 HTVDLTIEMMLDSQASDLD.....KLRTAKGRMGFLVAKKK 473
2573.449 Million cell updates/sec

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xih
-Q=/cgn2_1/USPTO_spool/US10031331/runat_26072004_120352_15000/app_query.fasta_1.647
-DB=Issued Patents NA -QFIX=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031331 cgn2_1.1.56 @runat_26072004_120352_15000 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	16.1	286	4	US-09-313-294A-7492
2	363	14.4	280	4	US-09-313-294A-6753
3	347	13.8	274	4	US-09-313-294A-3443
4	330	13.1	260	4	US-09-313-294A-876
5	262	10.4	174	4	US-09-313-294A-5689
6	257.5	10.2	174	4	US-09-313-294A-1188
7	218.5	8.7	852	4	US-09-313-294A-1188
8	218.5	8.7	852	4	US-09-313-294A-1188
9	210.5	8.4	954	4	US-09-313-294A-1188
10	201.5	8.0	6085	3	US-09-029-603-4
11	194.5	7.7	80161	3	US-09-036-987A-1
12	194.5	7.7	80161	3	US-09-370-700-1

C	13	194.5	7.7	80161	4	US-09-603-207-1	Sequence 1, Appli
	14	185.5	7.4	2122	3	US-09-029-603-1	Sequence 1, Appli
	15	184.5	7.3	750	4	US-09-134-000C-3308	Sequence 3308, Ap
	16	184.5	7.3	957	4	US-09-382-906A-1	Sequence 1, Appli
	17	181	7.2	828	4	US-09-266-965-35	Sequence 35, Appli
	18	175	7.0	1790	4	US-09-118-637A-3	Sequence 3, Appli
	19	174.5	6.9	1080	4	US-09-328-352-3879	Sequence 3879, Ap
	20	172	6.8	80161	3	US-09-036-987A-1	Sequence 1, Appli
	21	172	6.8	80161	3	US-09-370-700-1	Sequence 1, Appli
	22	172	6.8	80161	4	US-09-603-207-1	Sequence 1, Appli
C	23	167.5	6.7	997	4	US-08-956-171B-981	Sequence 981, App
	24	167.5	6.7	1369	4	US-09-439-554-29	Sequence 29, Appli
	25	165	6.6	771	4	US-09-328-352-3576	Sequence 3576, Ap
	26	163	6.5	1149	2	US-08-844-305-1	Sequence 1, Appli
	27	162	6.4	11466	4	US-08-956-171B-444	Sequence 444, App
C	28	158	6.3	984	4	US-09-252-991A-12173	Sequence 12173, A
	29	158	6.3	1254	4	US-09-252-991A-12420	Sequence 12420, A
C	30	156.5	6.2	777	4	US-09-134-001C-407	Sequence 407, App
	31	156	6.2	1400	3	US-09-041-718-1	Sequence 1, Appli
	32	155.5	6.2	1314	4	US-09-644-907B-1	Sequence 1, Appli
	33	154.5	6.1	888	4	US-09-540-236-1514	Sequence 1514, Ap
	34	154.5	6.1	954	4	US-08-489-039A-873	Sequence 873, App
	35	154.5	6.1	82909	4	US-09-536-002-32	Sequence 32, Appli
	36	153.5	6.1	22108	3	US-09-085-761A-3	Sequence 3, Appli
	37	153.5	6.1	22108	4	US-09-085-761A-3	Sequence 3, Appli
C	38	152	6.0	4403765	3	US-09-103-840A-2	Sequence 1, Appli
C	39	152	6.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	40	151	6.0	1173	4	US-09-489-039A-2534	Sequence 2534, Ap
	41	150	6.0	171	4	US-09-313-294A-3925	Sequence 3925, Ap
	42	149	5.9	762	4	US-09-489-039A-913	Sequence 913, App
	43	149	5.9	783	4	US-09-252-991A-12098	Sequence 12098, A
C	44	147.5	5.9	1131	4	US-09-252-991A-2322	Sequence 2322, Ap
	45	147.5	5.9	1215	4	US-09-252-991A-2513	Sequence 2513, Ap

ALIGNMENTS

RESULT 1

US-09-313-294A-7492
; Sequence 7492, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalngudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 7492
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700381824H1
US-09-313-294A-7492

Alignment Scores:

Prod. No.: 1.65e-43
Score: 405.00
Percent Similarity: 84.21%
Best Local Similarity: 75.79%
Query Match: 16.10%
DB: 4
Length: 286
Matches: 72
Conservative: 8
Mismatch: 15
Indels: 0
Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-7492 (1-286)

QY 134 GlyTyrIlePhePheArgGluSerCysPheHISGLNserGlyAspHisLysArgLysSer 153
DB 2 GGCCTATTTTCTTTAGAGNATCATGTTTCCACCATCTGGAGATTCCTCAAAAGGAGTG 61

Qy	154	AsnProThrHieTyArgGluProArgPheTyrThrLysAlaPheLysGluCyHisLeu	173
Db	62	AACCCAACACACTATCGAGAACCAAGGTTTTATACCACAGGTATTTAAAGAGGGCCATTCA	121
Qy	174	GlnAspGlySerClyAasnSerTyrGluLeuSerLeuSerCysLysCysValLeuGlyAla	193
Db	122	TTTGATCAAGTGGAGGTTGGTTGAACTTCTCTAGTCACCTGTAAATGATATGGGGCT	181
Qy	194	TyrValArgAsnLysLysaenGlnaenGlnIleSerTyrIleuTyrGlnLysValaAspSer	213
Db	182	TATGTCAAAACAAGAGGAATCAAAACCAATATGCTGCTTATGGGAAGAAGGTAAATCA	241
Qy	214	LysAspAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyr	228
Db	242	ACAGACACGACGATTTTCAAGATTTCTGGACAACTGTCATAC	286

RESULT 2

```

US-09-313-294A-6753
; Sequence 6753, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V..
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1993-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6753

```

US-10-031-331B-40 (1-473) x US-09-313-294A-6753 (1-280)

QY	359	ProGluGlyIysValIleuIleSerAspTyrCysValIysAlaGlyProProSerProGlu	378
Db	2	CCTGGGGCAGGTCTCTTATCAGTGACTACTCTGAGGAGTCTCGGAAACCATCAGAGAA	61
QY	379	PhaAlaIaIaTzIleIysGlnArgGlyTyrAspLeuHisAspValIysGluTyrGlyGln	398
Db	62	TTTCGACGGTACATTAAAGCAGAGAGTTATGACCTACATGCTGTGGAGCTTATGGACAG	121
QY	399	MetIeuIysAspAlaGlyPheValAspValIeuAlaGluAspArgThrGluGlnPheIle	418
Db	122	NGTTGAAGAGTCTGGTTTCGTGTGTCATTTCTGAGGATCGAATCGATCAGTCTCTT	181
QY	419	ArgValIeuArgIysGluGluGlnIuThrValGluIysGluIysAspValPheIleSerAsp	438
Db	182	GGTGTTTTAGTAAGAGTTAGTCTGCAATTTGAAAGACACAGGACGNTTCTCTCTGAC	241
QY	439	PheSerGluGluAspTyrAsnAspIleValGlyGlyTrp	451
Db	242	TTTACCCAGGAGGACTACGATGATATCGTGACGGTTGG	280

RESULT 3

US-09-313-294A-3443
; Sequence 3443, Application US/09313294A

```

; Patent No. 6476212
;
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3443
; LENGTH: 274
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID NO. 6476212 700611594H1
; NAME/KEY: unsure
; LOCATION: 92, 202
; OTHER INFORMATION: a, t, c, g, or other
; US-09-313-294A-3443
;
; Alignment Scores:
; Pred. No.: 5,52e-36
; Score: 247.00
; Percent Similarity: 85.06%
; Best Local Similarity: 78.16%
; Query Match: 13.80%
; DB: 4
; Indels: 1
; Gaps: 0
; Conservatives: 6
; Matches: 68
; Length: 274

```

US-10-031-331B-40 (1-473) x US-09-313-294A-3443 (1-274)

Qy	219	PheGlnAaGpeLeuAspThrSerClnrTyrlvsCysaasSerileLeuargTyrcGuarg	238
		
Db	5	TTTCAAGATTCCTGGACACGTCGATACAAACAGTCGGATATTAGTACAGCGCT	64
		
Qy	239	valPheGlyproGlyTyrrValSerThrGlyGlyTyrcGluThrThryGluPheValSer	258
		
Db	65	GTCCTTTGGTGAAGGTTTGTGACACNCTGTGGATCTCGACACAAAGAGGATTTGTGGC	124
		
Qy	259	MetLeuAspLeuLysproGlyGlyLysValLeuAspValGlyCysGlyIleGlyGly	278
		
Db	125	ATGCTCGATCTTAAACCGGGCCAGAAAGTACTTGATCTCGATGTGGATTTGGAGCGCG	184
		
Qy	279	AspPheTyrrMetAlaGluThrPheAspVal-GluValValGlyPheAspLeuSerValAs	298
		
Db	185	GACTTTTACATCGCTGCNAACCTATGATGTCCCATGTTCTTGGTATTGATCTATCGGTGAA	244
		
Qy	298	nMetIleSerPheAlaLeu	304
		
Db	245	CATGGTTTTCATTTCGAATG	263

PRISIT.T 4

```

RESULTS 4
US-09-313-294A-876
/ Sequence 876, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Lalgudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 876
/ LENGTH: 260
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature

```

OTHER INFORMATION: Incyte ID No. 6476212 700549985H1
 NAME/KEY: unsure
 LOCATION: 24, 86
 OTHER INFORMATION: a, t, c, g, or other
 US-09-313-294A-876

Alignment Scores:
 Pred. No.: 8 29e-34 Length: 260
 Score: 330.00 Matches: 66
 Percent Similarity: 83.53% Conservative: 5
 Best Local Similarity: 77.65% Mismatches: 12
 Query Match: 13.12% Indels: 2
 DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-876 (1-260)

QY 201 GlnGlnIleSerThrLeu-TpGlnIleValAspSerIleLeuArgValPheGlu 220
 Db 3 CAAACACGATATCTGCTTANTGGGAAGGTAAATACACAGACAGAGATTTTCA 62
 QY 220 nArpPheLeuAsp-ThrSerGlnIleValLeuArgValPheGluArgValP 240
 Db 63 AGATTCCTGACACACCTGCAATNCCAAACCAAGTGGGATATTACGTTACGCGTGTCT 122
 QY 240 heGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLeuValSerMetL 260
 Db 123 TTGGTGAAGGTTTGTGAGCACTGGTGGATCGACACACAAAGGATTTGTGGCGATGC 182
 QY 260 euAspLeuIleProGlyGlnIleValLeuAspValGlyCysGlyIleGlyGlyAsp 280
 Db 183 TCGATCTTAACCGCGCCAGAAAGTACTTGTATGTGGATGTGGAAATGGAGCGCGCACT 242
 QY 280 heTyrMetAla 283
 Db 243 TTTACATGGCT 253

RESULT 5

US-09-313-294A-5689
 Sequence 5689, Application US/09313294A
 Patent No. 6476212

GENERAL INFORMATION:
 APPLICANT: Lalgudi, Raghunath V.
 APPLICANT: Ito, Laura Y.
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
 FILE REFERENCE: PL-0017 US
 CURRENT APPLICATION NUMBER: US/09/313,294A
 CURRENT FILING DATE: 1999-05-14
 NUMBER OF SEQ ID NOS: 7600
 SOFTWARE: PERL Program
 SEQ ID NO 5689
 LENGTH: 174
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. 6476212 700350535H1
 US-09-313-294A-5689

Alignment Scores:
 Pred. No.: 3 03e-25 Length: 174
 Score: 262.00 Matches: 49
 Percent Similarity: 89.66% Conservative: 3
 Best Local Similarity: 84.48% Mismatches: 6
 Query Match: 10.42% Indels: 0
 DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-5689 (1-174)

QY 234 LeuArgTyrGluArgValPheGlyProGlyTyrValSerThrGlyValTyrGluThr 253
 Db 1 TTACGTTACGAGCGTGTCTTTGTGGAAGGTTTGTGAGCACTGGTGGAAATCGACACACA 60

QY 254 LysGluPheValSerMetLeuAspLeuLysProGlyGlnIleValLeuAspValGlyCys 273
 Db 61 AAGGAATTTGGGCGCATCTCGATCTTAACCGCGCCAGAAAGTACTTGTATGTGGATGT 120
 QY 274 GlyIleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValVal 291
 Db 121 GGAATTGGAGCGCGGACCTTTTACATGGCTGCAAACTATGATGTCCATGTTCTT 174

RESULT 6

US-09-313-294A-1188
 Sequence 1188, Application US/09313294A
 Patent No. 6476212

GENERAL INFORMATION:
 APPLICANT: Lalgudi, Raghunath V.
 APPLICANT: Ito, Laura Y.
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
 FILE REFERENCE: PL-0017 US
 CURRENT APPLICATION NUMBER: US/09/313,294A
 CURRENT FILING DATE: 1999-05-14
 NUMBER OF SEQ ID NOS: 7600
 SOFTWARE: PERL Program
 SEQ ID NO 1188
 LENGTH: 274
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. 6476212 700550521H1
 NAME/KEY: unsure
 LOCATION: 256
 OTHER INFORMATION: a, t, c, g, or other
 US-09-313-294A-1188

Alignment Scores:

Pred. No.: 2 53e-24 Length: 274
 Score: 257.50 Matches: 62
 Percent Similarity: 80.43% Conservative: 12
 Best Local Similarity: 67.39% Mismatches: 18
 Query Match: 10.24% Indels: 4
 DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x US-09-313-294A-1188 (1-274)

QY 48 ArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnValIleAlaLeuAspPheIleGlu 67
 Db 2 CGCTTTACTGGAGTCTGGCAAGAAGCTGGGCGACGTTCTGGCTCTAGACTTCATTGAA 61
 QY 68 SerAlaIleLysLysAsnGluValIleAsnGlyHisTyrLysAsnValLysPheMetCys 87
 Db 62 AGTGTGATTGAAGAACAACCAAGCATATATGGGCATCACAGACACATACCTTCAGGTGT 121
 QY 88 AlaAspValThrSerProThrLeuSerPheProHisSerLeuAspValIlePheSer 107
 Db 122 GCTGATGTGACATCTAACGACTTGAAGATTGAAT-AACTCTGTGATCTGATATTTTCA 180
 QY 108 AsnTrpLeuLeuMetTyrLeuSerAspGluGluValGluAsnLeuValGluArgMetLeu 127
 Db 181 AACTGGCTATTATATGTACCT-TCAGATGGAGGAGGTCCAAAGCTTGTGGG-GAAATGTGT- 237
 QY 128 LysTrpLeuLysProGlyGlyTyrIlePhePheArg 139
 Db 238 AATGGCTAATAGTCGGGNGCCATATTTCTTTTGA 273

RESULT 7

US-09-266-965-36
 Sequence 36, Application US/09266965
 Patent No. 6495348

GENERAL INFORMATION:
 APPLICANT: Sherman, D
 APPLICANT: Mao, Y
 APPLICANT: Varoglu, M
 APPLICANT: He, M

; APPLICANT: Sheldon, P
 ; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
 ; FILE REFERENCE: 600.456US1
 ; CURRENT APPLICATION NUMBER: US/09/266,965
 ; EARLIER FILING DATE: 1999-03-12
 ; EARLIER FILING DATE: 1996-08-19
 ; EARLIER APPLICATION NUMBER: PCT/US94/11279
 ; EARLIER FILING DATE: 1994-10-06
 ; EARLIER APPLICATION NUMBER: US 08/133,963
 ; EARLIER FILING DATE: 1993-10-07
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 36
 ; TYPE: DNA
 ; LENGTH: 852
 ; ORGANISM: Streptomyces lavendulae
 ; US-09-266-965-36

Alignment Scores:
 Pred. No.: 2,12e-18 Length: 852
 Score: 218.50 Matches: 68
 Percent Similarity: 45.06% Conservative: 46
 Best Local Similarity: 26.88% Mismatches: 92
 Query Match: 8.69% Indels: 47
 DB: 4 Caps: 11

US-10-031-331B-40 (1-473) x US-09-266-965-36 (1-852)

```

Qy 236 TyrGluArgValPheGlyProGlyTyrValSerThrGly-----GlyTyr 250
Db 67 TAGGACCGGTTACCGCGCTGGGAGCGCTCTCGGAGAACCTGCACCTTCGGCTAC 126

Qy 251 -----GlutThrThrysgluPheValSerMet 259
Db 127 TGGGACTCCCGCAGCAGCGTGGCGTGGCGGACCGCCAGCGGCTCACCGACATG 186

Qy 260 -----LeuAspLeuLeuPheGlyProGlyValLeuAspValGlyCysGlyIle 275
Db 187 ATGGCGGAGCGGCTCGCGTCTGGCGGCTCGCGGCTCGCGGCTCGCGGCTCGCGG 246

Qy 276 GlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeu 295
Db 247 GGGACCGCGGCGTACGCGTACGCGGCTCGCGGCTCGCGGCTCGCGGCTCGCGG 306

Qy 296 SerValAsnMetIle-----SerPheAlaLeuGluArgSerIleGlyLeuLysCysAla 313
Db 307 AGCGATGACGAGTGTCTCGGCGCCACGCGCTGGCGGAGCGCGGCTCGCGGACCG 366

Qy 314 ValGluPheGluValAlaAspCysThrIleAsnTyrProAspAsnSerPheAspVal 333
Db 367 GCGCGCTTCACGCGCGCGCGCGCGATGCGGCTCGCGGCTCGCGGCTCGCGGCTCG 426

Qy 334 IleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPhe 353
Db 427 GTCATCGCCCTCGATCGATCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGG 486

Qy 354 TyrLysTrpLeuLysProGlyGlyValLeuLeuSerAspTyrCysLysLysAlaGly 373
Db 487 GCGCGGCTGTCTCGCGCGCGCGCGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 543

Qy 374 ProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAsp--- 392
Db 544 CCCTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594

Qy 393 -----ValLysGluTyrGlyGlnMetLeuLysAspAlaGly----- 404
Db 595 ATGATGACCGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654

Qy 405 -----PheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeu 421
Db 655 CTGAGAGAGTTCCTCGACATC---AGCGACGACGACCGCTGGAGAGACCTTCAGG 711
  
```

```

Qy 422 ArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGlu 441
Db 712 TCG-----GAGCGCATCAACTCTCGAAGCAGAGAGCTGGAGACCGCACTTCGGCAG 762

Qy 442 Glu-----AspTyrAsnAspIleValGly 449
Db 763 GAGATGCTGAACCACTTCGACCCCGCGGACCTCGTCGCGC 801
  
```

RESULT 8

US-09-266-965-76/c
 ; Sequence 76, Application US/09266965
 ; Patent No. 6495348
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D
 ; APPLICANT: Mao, Y
 ; APPLICANT: Varoglu, M
 ; APPLICANT: He, M
 ; APPLICANT: Sheldon, P
 ; TITLE OF INVENTION: Mitomycin biosynthetic gene cluster
 ; FILE REFERENCE: 600.456US1
 ; CURRENT APPLICATION NUMBER: US/09/266,965
 ; EARLIER FILING DATE: 1999-03-12
 ; EARLIER FILING DATE: 1996-08-19
 ; EARLIER APPLICATION NUMBER: PCT/US94/11279
 ; EARLIER FILING DATE: 1994-10-06
 ; EARLIER APPLICATION NUMBER: US 08/133,963
 ; EARLIER FILING DATE: 1993-10-07
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 76
 ; TYPE: DNA
 ; LENGTH: 53500
 ; ORGANISM: Streptomyces lavendulae
 ; US-09-266-965-76

Alignment Scores:
 Pred. No.: 2,46e-15 Length: 53500
 Score: 218.50 Matches: 68
 Percent Similarity: 45.06% Conservative: 46
 Best Local Similarity: 26.88% Mismatches: 92
 Query Match: 8.69% Indels: 47
 DB: 4 Caps: 11

US-10-031-331B-40 (1-473) x US-09-266-965-76 (1-53500)

```

Qy 236 TyrGluArgValPheGlyProGlyTyrValSerThrGly-----GlyTyr 250
Db 7518 TAGGACCGGTTACCGCGCTGGGAGCGCTCTCGGAGAACCTGCACCTTCGGCTAC 7459

Qy 251 -----GlutThrThrysgluPheValSerMet 259
Db 7458 TGGGACTCCCGCAGCAGCGTGGCGTGGCGGACCGCCAGCGGCTCACCGACATG 7399

Qy 260 -----LeuAspLeuLeuPheGlyProGlyValLeuAspValGlyCysGlyIle 275
Db 7398 ATGGCGGAGCGGCTCGCGTACGCGGCTCGCGGCTCGCGGCTCGCGGCTCGCGG 7339

Qy 276 GlyGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeu 295
Db 7338 GGGACCGCGGCTGAGCATGCGCGGCTCGCGGCTCGCGGCTCGCGGCTCGCGG 7279

Qy 296 SerValAsnMetIle-----SerPheAlaLeuGluArgSerIleGlyLeuLysCysAla 313
Db 7278 AGCCATGACGAGTGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7219

Qy 314 ValGluPheGluValAlaAspCysThrIleAsnTyrProAspAsnSerPheAspVal 333
Db 7218 GCGCGCTTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7159

Qy 334 IleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPhe 353
Db 7158 GTCATCGCCCTCGATCGATCATCCACATCGCGCGCGCGCGCGCGCGCGCGCG 7099
  
```

QY 354 TyrLeuTrpLeuLysProGlyGlyValLeuLeuLeuSerAspTyrCysValValAlaGly 373
Db 7098 GCGCGGTGTGGCGCGCGCGCGCTGTGGTCTCAGCGCTCTTTCGAGCGCGCC--- 7042
QY 374 ProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAsp--- 392
Db 7041 CCCCTCGCGCGCGCGCGCGCGCGCTGTGGTCTCAGCGCTCTTTCGAGCGCGCC--- 6991
QY 393 -----ValLysGluTyrGlyGlnMetLeuLysAspAlaGly----- 404
Db 6990 ATGATGACCATGTCAGCGCGCGCGCGCTGTGGTCTCAGCGCTCTTTCGAGCGCGCC--- 6931
QY 405 -----PheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeu 421
Db 6930 CTGGAGGAGTTCTCGCATC---AGCGACGACGCTCGAGAGAGCTTCAGGCTGCTC 6874
QY 422 ArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspHisSerGlu 441
Db 6873 TCG-----GAGCGCATCAACTCTCGAGCAGAGGCTCGAGACGACGCTTCGCGAG 6823
QY 442 Glu-----AspTyrAsnAspIleValGly 449
Db 6822 GAGATGTCGACCATGTCGACCGCGCGCGCTGTGGTCTCAGCGCTCTTTCGAGCGCGCC--- 6784

RESULT 9
US-09-118-637A-1
; Sequence 1, Application US/09118637A
; Patent No. 642434
; GENERAL INFORMATION:
; APPLICANT: Dellapenna, Dean
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH TOCOPHEROL
; TITLE OF INVENTION: METHYLTRANSFERASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/118, 637A
; APPLICATION NUMBER: US/09/118, 637A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 920905.90024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 954 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..954
US-09-118-637A-1

Alignment Scores:
Pred. No.: 2.83e-17 Length: 954
Score: 210.50 Matches: 55

Percent Similarity: 49.52% Conservative: 48
Best Local Similarity: 26.44% Mismatches: 72
Query Match: 8.37% Indels: 33
DB: 4 Gaps: 7
US-10-031-331B-40 (1-473) x US-09-118-637A-1 (1-954)
QY 236 TyrGluArgValPheGlyPro-----GlyTyrValSerThrGlyGlyTyrGlu 251
Db 145 TCGGAAGACGTTTGGGTGAGCATATGCAACCGCTACTAGCTCCCGCAGCCTAT 204
QY 252 ThrThrLysGluPheValSerMetLeuAspLeu----- 262
Db 205 CGGATCGATCGCGCCAGGCTCAATTTGATCTGATCAAGAAGTAATTGGCTGGCGAGTG 264
QY 263 -----LysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGly 277
Db 265 CCCCAAAATAGCGCAACCA---CGAAAAATTTCTCGATTTAGGCTGTGGCATTTGGCGC 321
QY 278 GlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerVal 297
Db 322 AGTAGTTTGTACTTGGCCAGCAACCAAGCAGAGTAAGTGTATGGGGCTAGTCTTTCCCA 381
QY 298 AsnMetIleSerPheAlaLeuGlu-----ArgSerIleGlyLeuLysCysAlaValGlu 315
Db 382 GTGAGGTGGAGCGCGCGGGAAGGGCCAGGCGCTGGGGTTGGGCTCAACCTGGCGAG 441
QY 316 PheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyr 335
Db 442 TTTGAGGTGGCAATGCTTTGGGATTTTGGCTTCCGATTTCTTTCAGCTGGGTTGG 501
QY 336 SerArgAspThrIleLeuHisIleGlnAspValProAlaLeuPheAspSerPheTyrLys 355
Db 502 TCGTTGGAAGTGGGGAGCAGCATGCCCAACAAAGCTCAGTTTTCACAGAAGCTTGGCGG 561
QY 356 TrpLeuLysProGlyGlyLysValLeuLeuSerAspTyrCysLysLysAla----- 372
Db 562 GTACTTAACACGAGTGGCGCTGTGATTTAGCGACCTGTGTCTCATCTGCCCATTCATCCC 621
QY 373 -----GlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu 390
Db 622 GCGAATGCGCGCGCTGACTGCGCATGAACGTGGCCATCTC---CAAGCCATCTATGACGT 678
QY 391 His-----AspValLysGluTyrGlyGlnMetLeuLysAspAlaGly 404
Db 679 TACTGTTGGCTTATGTGTTTCCCTCGGACTACGAGCGATCCCGAGGATGTGGG 738
QY 405 PheValAspValLeuAlaGluAsp 412
Db 739 TTTGGGAATTAAGACTGCGCAT 762

RESULT 10
US-09-029-603-4
; Sequence 4, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospiech, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029, 603
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6085
; TYPE: DNA
; ORGANISM: Streptomyces longisporoflavus

QY 247 ThrGlyGlyTyr-----GluThr 252
Db 13487 ---GGGGGATCGGAGGATCGGGTGGCCACACCGTGTGGATCGTGGCGACACTG 13431
QY 253 ThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGly 272
Db 13430 ACCGACCTGTTTCATCGCAAGCGCGCGCTCGGTCCGGAGCGCACCTGTTTCGACTGGGC 13371
QY 273 CysGlyIleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGly 292
Db 13370 TGGCGCAATGGCGACCGCGTGTGGTGGCGATCGCGCGCGGTTCGAGTCACCGGA 13311
QY 293 PheAspLeuSerValLeuMetIleSerPheAla-----LeuGluArgSerIleGlyLeu 310
Db 13310 ATCACCCTGACCGCCAGCATCTGCGCGCGCCACACCGCTCGCCACGACGCGACTG 13251
QY 311 LysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSer 330
Db 13250 GCGGAGTCTTGTAGTTTCATCTAGTCGACGCGCGCGCGCTGCTCCCGGAGCATCATCACT 13191
QY 331 PheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPhe 350
Db 13190 TTTACGCGCGCATGGCGATCGCTGCTGTCAGATCGTGACCGCGCGCGATC 13131
QY 351 ArgSerPheTyrLysTyrLeuLysProGlyGlnLysValLeuLysAspTyrCysLys 370
Db 13130 CCGGAGTCCACCGAATCTCGAACCCTGGACCGCGCGCGGTTCGCTCCGAGACATCATCACT 13071
QY 371 LysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu 390
Db 13070 CCGGTTTCGACTCCCGGAA---GAGTACGCGCGCGTT-----TGGACGCGCGACCGCC 13020
QY 391 HisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAla 410
Db 13019 CATACCTTGACACTTCACGGCGCTGGTCAGGAGCGCGGTTCGAGATCTCGAGTC 12960
QY 411 GluAspArgThrGluGln-----PheIleArgValLeuArgLys 423
Db 12959 ACCGACCTCAGCGACGACGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 12900
QY 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443
Db 12839 TTGGGAGACATC 12828

RESULT 12

US-09-370-700-1/c
; Sequence 1. Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Tredey, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

Alignment Scores: 6.56e-12 Length: 80161
Pred. No.: 194.50 Matches: 74
Score: 41.67% Conservatives: 61
Best Local Similarity: 22.84% Mismatches: 137
Query Match: 7.73% Indels: 53
DB: 3 Gaps: 11
US-10-031-331B-40 (1-473) x US-09-370-700-1 (1-80161)
QY 172 HisLeuGlnAspGlySerGlyAsnSerTyrGluLeuSerLeuSerCysLysCysIle 191
Db 13786 CACGAGGAGCGCGCATCGGGCTGTACGAGCGGAGTACTTGGAAAGCAATTTGTC 13727
QY 192 GlyAlaTyrValArgAsn-----LysLysAsnGlnAsnGln-----Ser 213
Db 13726 GCGCTCCACACCTACCAACACATCGCGTTCCTGGAGAAAGCGTCAACGCCGAGGCGC 13667
QY 204 ---IleSerTyrLeuTyrGlnLysValAsp-----Ser 213
Db 13666 GTTCTGCTTGGTGGCCAGGAGTCTGGAGCATATTCACCTGGCGCGTGAACGAC 13607
QY 214 LysAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyr-----Lys 229
Db 13606 CCGGAGGAGCAGTG--AACAGCAGAGCGCGGACACACAGGCAATTTCCGACCGCGGATCAG 13548
QY 230 CysAsnSerIle-----LeuArgTyrGluArgValPheGlyProGlyTyrValSer 246
Db 13547 GTGAGTCTCATTTGATGCGTTGGCGCGCGGTCCCTGCACACGGTTACTGGCG 13488
QY 247 ThrGlyGlyTyr-----GluThr 252
Db 13487 ---GGCGGTATCGGAGGATCGCGTGGCCACACCGTGTGGATGTGCGACCACTG 13431
QY 253 ThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGly 272
Db 13430 ACCGACCTGTTTCATCGCAAGCGCGCGCTCGGTCCGGAGCGCACCTGTTTCGACTGGGC 13371
QY 273 CysGlyIleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGly 292
Db 13370 TGGCGCAATGGCGACCGCGTGTGGTGGCGATCGCGCGCGGTTCGAGTCACCGGA 13311
QY 293 PheAspLeuSerValLeuMetIleSerPheAla-----LeuGluArgSerIleGlyLeu 310
Db 13310 ATCACCCTGACCGCCAGCATCTCGCGCGCGCCACACCGCTCGCCACGACGCGGACTG 13251
QY 311 LysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSer 330
Db 13250 GCGGAGTCTTGTAGTTTCATCTAGTCGACGCGCGCGCGGTTCGAGATCTCGAGTC 13191
QY 331 PheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPhe 350
Db 13190 TTTACGCGCGCATGGCGATCGCTGCTGTCAGATCGTGACCGCGCGCGATC 13131
QY 351 ArgSerPheTyrLysTyrLeuLysProGlyGlnLysValLeuLysAspTyrCysLys 370
Db 13130 CCGGAGTCCACCGAATCTCGAACCCTGGACCGCGCGCGGTTCGCTCCGAGACATCATCACT 13071
QY 371 LysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu 390
Db 13070 CCGGTTTCGACTCCCGGAA---GAGTACGCGCGCGTT-----TGGACGCGCGACCGCC 13020
QY 391 HisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAla 410
Db 13019 CATACCTTGACACTTCACGGCGCTGGTCAGGAGCGCGGTTCGAGATCTCGAGTC 12960
QY 411 GluAspArgThrGluGln-----PheIleArgValLeuArgLys 423
Db 12959 ACCGACCTCAGCGACGACGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 12900
QY 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443
Db 12839 TTGGGAGACATC 12828

Db 12899 AACTCGATGAGCTGCGCGGCTGACAGCTGTCGGGCTGTGCGCACCTTACACAGCAACGGTAC 12840
 QY 444 TyrAsnAspIle 447
 Db 12839 TTGGGAGACATC 12828
 RESULT 13
 US-09-603-207-1/c
 ; Sequence 1, Application US/09603207B
 ; Patent No. 6521406
 ; GENERAL INFORMATION:
 ; APPLICANT: Baltz, Richard H
 ; APPLICANT: Broughton, Mary C
 ; APPLICANT: Crawford, Kathryn P
 ; APPLICANT: Madduri, Krishnamurthy
 ; APPLICANT: Treadway, Patti J
 ; APPLICANT: Turner, Jan R
 ; APPLICANT: Waldron, Clive
 ; TITLE OF INVENTION: Bicosynthetic Genes For Spinosyn Insecticide
 ; FILE REFERENCE: 50489 DIV1
 ; CURRENT APPLICATION NUMBER: US/09/603,207B
 ; CURRENT FILING DATE: 2000-06-23
 ; EARLIER APPLICATION NUMBER: 09/370,700
 ; EARLIER FILING DATE: 1998-03-09
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 80161
 ; TYPE: DNA
 ; ORGANISM: Saccharopolyspora spinosa
 US-09-603-207-1

Alignment Scores:
 Pred. No.: 6.56e-12 Length: 80161
 Score: 194.50 Matches: 74
 Percent Similarity: 41.67% Conservative: 61
 Best Local Similarity: 22.84% Mismatches: 137
 Query Match: 7.73% Indels: 53
 DB: 4 Gaps: 11

US-10-031-331B-40 (1-473) x US-09-603-207-1 (1-80161)

QY 172 HsLeuGlnAspGlySerGlyAsnSerTyrGluLeuSerLeuLeuSerCysLysCysIle 191
 Db 13786 CACGAGGAGAGCCGCGATGCGGGCTGTGACGCGGAGCTACCTCGAAGCGCAATTGGTTC 13727
 QY 192 GlyAlaTyrValArgAsn-----LysLysAsnGlnAsnGln----- 203
 Db 13726 GGCCTCCACCTACCAACATCATCGCTTCTCGGAGAGGGCTCAACCGCGAAGCGGCGC 13667
 QY 204 ----IleSerTyrLeuTyrGlnLysValAsp-----Ser 213
 Db 13666 GTTCTGCTTGGTGGCCGAGGAGTCTGGACGACATATGTCACCTGGCCGACGTGAACAGC 13607
 QY 214 LysAspAspLysGlyPheGlnArgPheLeuAspThrSerGlnTyr-----Lys 229
 Db 13606 GCGGAGGACGAGTGAACAGCAGAGGGCGGCAACACAGGCAATTTCCGACCGCGGATCAG 13548
 QY 230 CysAsnSerIle-----LeuArgTyrGluArgValPheGlyProGlyTyrValSer 246
 Db 13547 GTGAGTTCATCTTCGATGCGTTGGCGACGGGGCTCCCTGACACCGATTACTGGGGC 13488
 QY 247 ThrGlyGlyTyr-----GluThr 252
 Db 13487 ---GGCGGTATCGGGAGATGCGGTCACACCGCTGTCGGATGCTGCCGACCACTG 13431
 QY 253 ThrLysGluPheValSerMetLeuAspLeuLysProGlyGlnLysValLeuAspValGly 272
 Db 13430 ACCGACCTGTTTCATGACAGCGCGCTCGTCCCGGAGCGGACCTGTTGACCTGGGC 13371
 QY 273 CysGlyIleGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValGly 292
 Db 13370 TGGCGCAATGGGACGCCCTGATCGCTCGCGGATGCGCCAGCGGCGTGTGAGTGCACCGGA 13311

QY 293 PheAspLeuSerValAsnMetIleSerPheAla-----LeuGluArgSerIleGlyLeu 310
 Db 13310 ATCACCGTGAACCCCGACGATCTCGCGCGCGCCACCGAGCTGCCACAGACGCGGACTG 13251
 QY 311 LysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSer 330
 Db 13250 GCGGCGAGTCTTGAGTTGATCTAGTCGACGCGCGCCAGCTGCCCTACCCGACGGTTTC 13191
 QY 331 PheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLysProAlaLeuPhe 350
 Db 13190 TTTGAGCCCGCATGGCGGATGTCAGTCCGTGTCGAGATCTGCGACCGCCGCCGATC 13131
 QY 351 ArgSerPheTyrLysTrpLeuLysProGlyGlyCysValLeuLysSerAspTyrCysLys 370
 Db 13130 CGCGAGGTCCACCGAATCTCGAACCAGCGCGCGCTGCTGCTCGGAGCATCATCACT 13071
 QY 371 LysAlaGlyProProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeu 390
 Db 13070 CGGGTTCGACTCCCGGAA---GAGTACGCGCGGTT-----TGACGCGGACGACGCC 13020
 QY 391 HisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAla 410
 Db 13019 CATACCTTGAACAGCTTCACGCGGCTGTCAGGAGCGCGGTTCGAGATTCTCGAATC 12960
 QY 411 GluAspArgThrGluGln-----PheIleArgValLeuArgLys 423
 Db 12959 ACCGACCTCACGCGCACACAGACAGCGTGCATGCTCTCTGCTAGCTCGACGAGTTGCTCGG 12900
 QY 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443
 Db 12899 AAATCGATGAGCTGCGCGGCTGACAGCTGTCGGGCTGTGCGCACCTTACACAGCTAC 12840

RESULT 14

US-09-029-603-1
 ; Sequence 1, Application US/09029603
 ; Patent No. 6210935
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Engel, Natalie
 ; APPLICANT: Bietenhader, Jurgen
 ; APPLICANT: Toupet, Christine
 ; APPLICANT: Pospiech, Andreas
 ; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
 ; FILE REFERENCE: 4-20555/A/PCT
 ; CURRENT APPLICATION NUMBER: US/09/029,603
 ; CURRENT FILING DATE: 1998-03-20
 ; EARLIER APPLICATION NUMBER: PCT/EP96/03643
 ; EARLIER FILING DATE: 1996-08-19
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2122
 ; TYPE: DNA
 ; ORGANISM: Streptomyces longisporoflavus
 ; FEATURE:
 ; NAME/KEY: misc_RNA
 ; LOCATION: (1)..(2122)
 ; OTHER INFORMATION: product = 2.1 kb region
 US-09-029-603-1

Alignment Scores:

Pred. No.: 2e-13 Length: 2122
 Score: 185.50 Matches: 68
 Percent Similarity: 42.15% Conservative: 42
 Best Local Similarity: 26.05% Mismatches: 92
 Query Match: 7.38% Indels: 59
 DB: 3 Gaps: 13

US-10-031-331B-40 (1-473) x US-09-029-603-1 (1-2122)

```

QY 236 TyrGluArgVal-----PheGlyPro----- 242
DB 895 TACGACCGCTGAGCGCGATCAACGACGCGCTGTTCAACCCCAATGTGCACATC 954
QY 243 GlyTyrValSerThrGlyGlyTyrGluThrThrGlu----- 255
DB 955 GGTATTTGGACACCCCGGCTCGAGCCACCATCGAGAGCGCGATGACCGGCTCACC 1014
QY 256 -----PheValSerMetLeuAspLeuLeuPheProGlyGlnLysValLeuAspValGlyCys 273
DB 1015 GATGTGTTTCATCGAACGCGCTGACGCGTACCGCACCTCCACGCTCTCGACCTCGGCTGC 1074
QY 274 GlyLeuGlyGlyAspPheTyrMetAlaGluThrPheAspValGluValValGlyPhe 293
DB 1075 GGGTGGCGCGCGCGCTCAGGCTGTCGGCGCACCGGGCAGCGGTACCGCGATC 1134
QY 294 AspLeuSerValanMetIleSerPheAla-----LeuGluArgSerIleGlyLeuLys 311
DB 1135 AGCATCAGCAGGAGCAGATCAGACCGCCCAACCGCTGGCCCGAGCGCGGCTGCC 1194
QY 312 CysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyrProAspAsnSerPhe 331
DB 1195 GACCGTCCGCTGTTCCAGCATGGCAGCGCATGAACCTGCCCTCGCCGACGCGCTGCTC 1254
QY 332 AspValIleTyrSerArgAspThrIleLeuHisIleGlaAspLysProAlaLeuPheArg 351
DB 1255 GACCGCTGATGGCGCTGAGTCACTGCCACATCCCGCAGCGGAGCGAGTGTTCACC 1314
QY 352 SerPheTyrLysTrpLeuLysProGlyGlyValLeuLysLeuIleSerAspTyrCysLysLys 371
DB 1315 GAGGTGTCGGGTGCTGCGCCCGGGCGGATGCTCTCACCACATCTTCGAGCGC 1374
QY 372 AlaGlyProProSerProGluPheAlaAlaTyrIleLysGluArgGlyTyrAsp----- 389
DB 1375 -----CACCGCGCGAAGCG-----GTACGACCGCGGATCGACAGTTC 1416
QY 390 -----LeuHisAspValLysGluTyrGlyGlnMetLeuLysAsp 402
DB 1417 TGCGCGACCTGATGTCACCGCGGACATCGACGACTAGTGGCGTGTGCGACCGC 1476
QY 403 AlaGly-----PheValAspValLeuAlaGluAspArgThrGluGlnPhe 417
DB 1477 TCCGGCTGCGCTGCGCGAGATGTCGAGTCT-----ACCGACGACGACC 1521
QY 418 IleArgValLeuArgLysGlu-----LeuGluThrValGluLysGluLysAspVal 434
DB 1522 ACGTGGCTTCGCGGACGAGATCGGAGGCTCGCGCGTGCAG-----GAGCGCCCGGTG 1578
QY 435 PheIleSerAsp-----PheSerGluGluAspTyr-----AsnAspIleVal 448
DB 1579 OCCATGACGAGGCGCACTTGGCTTCGGCGACGACTCCTTCAAGCGCGTCCGACCTGGCG 1638
QY 449 Gly 449
DB 1639 GGC 1641

```

RESULT 15

```

; Sequence 3308, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 3308
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-3308

```

```

Alignment Scores:
Pred. No.: 4,59e-14 Length: 750
Score: 184.50 Matches: 72
Percent Similarity: 43.06% Conservative: 49
Best Local Similarity: 25.62% Mismatches: 109
Query Match: 7.34% Indels: 51
DB: 4 Gaps: 12

```

US-10-031-331B-40 (1-473) x US-09-134-000C-3308 (1-750)

```

QY 209 GlnLysValAspSerLysAspAspLys-----GlyPheGlnArgPheLeu 223
DB 4 AGGAGAGTTTCAGATGAAGAAAGAAACAATATGATGATAATATATTTTTCAAAAATACAGT 63
QY 224 AspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheGlyProGly 243
DB 64 CAATGAGTCCGCTCCAG-----AAAGGA 87
QY 244 TyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeuAspLeuLys 263
DB 88 CTGGCTGGTGGCGGAGAAATGGGAGACTTTGAAAGAGATGCTACCT-----GATTTTAAG 141
QY 264 ProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysPhePheTyrMetAla 283
DB 142 ---GCTAAGCGTGTCTGATTTAGCATCGGCTATGGCATGGACCTGTATATATATCCGATG 198
QY 284 GluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIleSerPheAla 303
DB 199 GAAACCGTGTCTCTCTAGTAGGTGTGTATTTCTCATATAATGCTCGAAGTAGCA 258
QY 304 LeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluAlaAspCysThrLys 323
DB 259 AAAGGAAAAACCCATTTTCCACAG-----ATTGAATATGATGCTGTGCCATAGAGAT 312
QY 324 IleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIle 343
DB 313 CTGGATTTCCAGAGGAGACCTTTGATGATATACTAAGTTCCTTGGCTTCATATATGTA 372
QY 344 GlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGlyGlyLysVal 363
DB 373 GCACGACTATGAGAAATTTTAATAAAAAAGATATATAGCATGCTGAAGGCTGGTGCATTTA 432
QY 364 LeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAlaAla 381
DB 433 GTTTTACA-----GTTGAACATCTCTGTTTCTCTCATGGA 471
QY 382 -----TyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGlu 395
DB 472 ACACAGACTGTTATATACGAAAGGAGAAATACTGCATTTCCCGGTGGACAATAT 531
QY 396 TyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGlu 415
DB 532 TATTATGAGGCGAAACGACAGCTATGTTTGTG-----GAAGAAAGGTGATCA 579
QY 416 GlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434
DB 580 AAATATCATAGACACTGACCATATCTTAATACACTGCTTCCAAATAGTTTTATATA 639
QY 435 -----PheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyTyrAsn 452
DB 640 AATCAGATTGTGGAGCCACAGCGCCAGAGACATGATGATATTT---CCGGGATGGCG 696
QY 453 AspLysLeuArgArgThrAlaLysGlyGluGlnArgGlyPheValLysLysLys 472
DB 697 GATCAATGCGACGCCCATGATG-----CTGATTGTTATCGCAAAA 738
QY 473 Lys 473

```


DB 739 AAG 741
 |||

Search completed: August 2, 2004, 04:40:06
Job time : 219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2004, 23:10:34 ; Search time 544 Seconds
(without alignments)
3693.746 Million cell updates/sec

Title: US-10-031-331b-40
Perfect score: 2515
Sequence: 1 HTVDLTIANMMLDSQASDLD.....KLRRATKGEQRWGLFVAKKK 473

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US10031331/runat_26072004_120351_14972/app_query.fasta_1.647
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOEFL=0
-LIST=45 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-MODE=LOCAL -OUTFWT=pcp -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=JUS10031331 -CGCN_1_470 -runat_26072004_120351_14972 -NCPU=6 -ICPU=3
-NO MAP -LARGESQ=0 -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:
1: geneseqn1980s:
2: geneseqn1990s:
3: geneseqn2000s:
4: geneseqn2001as:
5: geneseqn2001bs:
6: geneseqn2002as:
7: geneseqn2003as:
8: geneseqn2003bs:
9: geneseqn2003cs:
10: geneseqn2004s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2515	100.0	1602	AA74206	DNA encod
2	2204	87.6	2235	AA16797	Spinach P
3	1829.5	72.7	1962	AAH44262	Physcomit
4	1270	50.5	1120	AA16798	Spinach P
5	905	36.0	636	AC18332	DNA clone
6	828	32.9	605	ACL25710	DNA clone
7	804.5	32.0	667	ACL18333	DNA clone
8	785	31.2	611	ACL25711	DNA clone

C	9	769	30.6	777	8	ACL19405	DNA clone
C	10	767.5	30.5	575	8	ACL18347	DNA clone
C	11	751	29.9	566	8	ACL19398	DNA clone
C	12	747.5	29.7	650	8	ACL18342	DNA clone
C	13	732	29.1	622	8	ACL18325	DNA clone
C	14	719	28.6	487	9	AD82005	Arabidops
C	15	697	27.7	589	8	ACL25708	DNA clone
C	16	682.5	27.1	517	8	ACL18349	DNA clone
C	17	661	26.3	736	8	ACL18395	DNA clone
C	18	658	26.2	619	4	AAH44246	Physcomit
C	19	657.5	26.1	587	8	ACL18344	DNA clone
C	20	647	25.7	671	8	ACL19401	DNA clone
C	21	646	25.7	738	8	ACL18354	DNA clone
C	22	641	25.5	712	8	ACL18373	DNA clone
C	23	639	25.4	727	8	ACL18362	DNA clone
C	24	633	25.2	676	8	ACL18366	DNA clone
C	25	632	25.1	648	8	ACL19403	DNA clone
C	26	629	25.0	663	8	ACL19412	DNA clone
C	27	623	24.8	621	8	ACL18331	DNA clone
C	28	617	24.5	659	8	ACL18389	DNA clone
C	29	612	24.3	624	8	ACL19409	DNA clone
C	30	612	24.3	649	8	ACL19408	DNA clone
C	31	612	24.3	652	8	ACL19407	DNA clone
C	32	611	24.3	623	8	ACL18326	DNA clone
C	33	610	24.3	696	8	ACL18382	DNA clone
C	34	604	24.0	697	8	ACL18372	DNA clone
C	35	603	23.9	527	8	ACL18374	DNA clone
C	36	601.5	23.9	527	8	ACL18334	DNA clone
C	37	598.5	23.8	488	8	ACL18346	DNA clone
C	38	594	23.6	436	8	ACL18323	DNA clone
C	39	592	23.5	692	8	ACL18391	DNA clone
C	40	590	23.5	684	8	ACL18377	DNA clone
C	41	589	23.4	477	8	ACL25703	DNA clone
C	42	585.5	23.3	536	8	ACL18350	DNA clone
C	43	578	23.0	695	8	ACL18380	DNA clone
C	44	574	22.8	631	8	ACL18356	DNA clone
C	45	574	22.8	673	8	ACL18378	DNA clone

ALIGNMENTS

RESULT 1

AA74206
ID AAF74206 standard; DNA; 1602 BP.
AC AAF74206;
XX
DT 06-AUG-2003 (revised)
DT 02-MAY-2001 (first entry)
XX
DE DNA encoding environmental stress tolerant protein SEQ ID 39.
XX
KW Environmental stress resistance; salt; heat; desert; transgenic plant;
KW ds.
XX
OS Suaeda japonica.
XX
PN WO200106006-A1.
XX
PD 25-JAN-2001.
XX
PF 19-JUL-2000; 2000WO-JP004862.
XX
PR 19-JUL-1999; 99JP-00235910.
PR 24-MAR-2000; 2000JP-00085377.
XX
(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Yamada A, Ozeki Y, Saito T.
XX
DR WPI; 2001-147355/15.
DR P-PSDB; AAB80627.

PT Screening method to obtain DNA encoding environmental stress resistance
 PT factor, useful for producing transgenic plants resistant to environmental
 PT stress.

XX Claim 65: Page 123-125; 167pp; Japanese.

XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 -
 CC AAB80639, which impart environmental stress resistance. The invention
 CC relates to a method for identifying DNA encoding proteins imparting
 CC environmental stress resistance. The method comprises inserting cDNA from
 CC a library originating in a salt-resistant organism into a host cell,
 CC culturing the transformants under conditions in which the untransformed
 CC host does not grow well, and selecting for viable clones. The method is
 CC useful for obtaining DNA encoding environmental stress resistance
 CC factors. The DNA encoding proteins conferring environmental stress
 CC resistance, can be used in the production of plants resistant to
 CC environmental stress, which can be cultivated in unfavourable
 CC environments such as deserts, salt damaged ground, cold regions and the
 CC oceans. They can be used for increasing the area of land covered by green
 CC plants, and desert greening and afforestation, in order to counter the
 CC effects of the increase in atmospheric carbon dioxide concentration. PCR
 CC primers AAF74219 and AAF74220 are used in an example illustrating the
 CC method of the invention. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 1602 BP; 454 A; 279 C; 383 G; 486 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,3e-263 Length: 1602
 Score: 2515.00 Matches: 473
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-031-331B-40 (1-473) x AAF74206 (1-1602)

Qy 1 HsThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 Db 1 CACACCGTGTGATTAACCTTGAAGCTATGATGCTGATCTTCAAGCTTCTGATCTTGAC 60
 Qy 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysValLeu 40
 Db 61 AAGAAGAACGCTGCTGAGATCTTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTGG 120
 Qy 41 GluLeuGlyValGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60
 Db 121 GAACCTGGGCTGGTATTGGTGGTTTACGGTGAATGGCTGAGAAAGCTGGCCAGGTT 180
 Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
 Db 181 ATTGCTCTGGATTTCAATGAGAGTGTCTATCAAGAAAGATGAAGTAATCAATGGGCACTAC 240
 Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
 Db 241 AAAAATGTCAGTTTATGTCGTGATGTCATCTCCACTCTCCAGTTTCCACCAT 300
 Qy 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120
 Db 301 TCATTGGATGATATTCCTCAATGGTACTCATGATCTTCTGATGAGAGGTGGAA 360
 Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyLysTrpIlePhePheArgGlu 140
 Db 361 AATTGGTTGAAAGATGTTGAATGGTTGAAGCCAGGGGGTTACATTTTCTTCACAGAA 420
 Qy 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
 Db 421 TCTGTGTTCCATCTGGGATCACAACGCAAAAGCAATCCACCCACTACTCCGTGAA 480
 Qy 161 ProArgPheTyrThrIleAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 Db 481 CCTAGGTTCTACACTAAGGCCCTTCAAGAGTGTCAATTTGCAAGATGATGTCGAACTCT 540
 Qy 181 TyrGluLeuSerLeuLeuSerCysLysCysValIleGlyAlaTyrValArgAsnLysAsn 200

Db 541 TATGAGCTCTCCCTACTTACTGTCAAATGTATTGGAGCTTATGTGAGAAACAGAAAC 600
 Qy 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
 Db 601 CAGAACCAAGATTAGTTGGTGGTGGCAAAAGTTGATCTTAAGGATGATAAGGGGTTCAG 660
 Qy 221 ArgPheLeuAspThrSerGlnTrpLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
 Db 661 CATTTCTGGAACTACCCAGTCAAGTGTATAGCATCTCCGATATGAGCGTGTATT 720
 Qy 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrLysGluPheValSerMetLeu 260
 Db 721 GGCCTCGTATTATGTAGCTAGGATATGAACCAACCAAGAGCTTTGTGTCAATGCTG 780
 Qy 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyLysValAspPhe 280
 Db 781 GACTTGAAGCTGGCCAGAGAGTCTTGGATTTGGTGGAAATGGTGGAGGTGACTTT 840
 Qy 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
 Db 841 TACATGGCGGAGACCTTTGATGTGAGGTGTGGATTTGATCTCTCCTTAAATGATT 900
 Qy 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 Db 901 TCTTTGCTTGGCTTGGAGCGTCTTATTTGGCTTAAATGTGTGTTGGTGGAGTAGCAGAT 960
 Qy 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340
 Db 961 TGCACCAAGATAAATACCTGATTAATCTTTTGTGATGTCATCTATAGCGGTGACACCAT 1020
 Qy 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
 Db 1021 CTGCATATTCAGGACCAAGCCTGCTGTTAGATCTCTTCAAAATGGTTGAAGCCAGGA 1080
 Qy 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
 Db 1081 GGTAAAGTTCTAAATCAGTGATTTACTGCAAGAAAGCTGGTCCACCTCACTGAATTCGCC 1140
 Qy 381 AlaTyrIleLysGlnArgGlyCysAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
 Db 1141 GCTTACATTAGCAGAGGGGATGATGATCTCATGATGAAGGATATGGGCAGATGCTT 1200
 Qy 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
 Db 1201 AAAGATGCTGGATTTGTTGATGTTCTTGGCAGGATAGAACTGAGCAGTTCAATCGAGTT 1260
 Qy 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
 Db 1261 CTACGGAGGAACTAGAGACTGTTGAGAGGAAAGGATGTTTCATTAGTCAATTTCTCT 1320
 Qy 441 GluGluAspTyrAsnAspIleValGlyTyrAsnAspLysLeuArgThrAlaLys 460
 Db 1321 GAGGAGGATTAACATGACATTTGGAGGTGGAAATGATTAAGTTGGAGGAGTCCAGG 1380
 Qy 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473
 Db 1381 GGTGAGCAACGATGGGCTGTTGTTGGTCCCAAGAAAG 1419

RESULT 2

AAD16797 standard; DNA; 2235 BP.

ID AAD16797

XX AAD16797;

AC AAD16797;

XX 29-NOV-2001 (first entry)

DT 29-NOV-2001 (first entry)

XX Spinach PEAMT DNA.

DE Spinach PEAMT DNA.

XX Spinach: S-adenosyl-L-methionine; phosphoethanolamine N-methyltransferase;

KW PEAMT; cellular intermediate; phospho-dimethyl-ethanolamine; choline;

KW phosphomono-methyl-ethanolamine; phosphocholine; phosphatidylcholine;

KW glycine betaine; choline-O-sulphate; lipid content alteration;

osmotic stress tolerance; nutritional value; transgenic plant;
cryoprotectant; ds.

Spinacia oleracea.

Key Location/Qualifiers
CDS 254..1738
/*tag= a
/product= "Spinach PEAMT protein"

W0200168870-A2.

20-SBP-2001.

15-MAR-2001; 2001WO-US008352.

15-MAR-2000; 2000US-00525885.

(UYFL) UNIV FLORIDA.

(UYCA-) UNIV CARNEGIE MELLON.

Hanson AD, Nuccio ML, Henry SA;

WPI; 2001-565796/63.

P-PSDB; AAB09760.

New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase polypeptides; useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine and for altering the lipid content in plants cells.

Claim 18; Page 108-109; 158pp; English.

The present sequence is a DNA encoding spinach S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase (PEAMT). The PEAMT sequences are useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine, phosphomono-methylethanolamine, choline, betaine. They are useful for altering the lipid content in plant cells. The polynucleotides are also useful for improving the osmotic stress tolerance of a plant and increasing the cryoprotectant properties of a plant. The present invention also relates to methods and compositions comprising PEAMT used for generating transgenic plants with increased nutritional value

Sequence 2235 BP; 623 A; 399 C; 516 G; 697 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1-61e-229	Length:	2235
Score:	2204.00	Matches:	408
Percent Similarity:	93.64%	Conservative:	34
Best Local Similarity:	86.44%	Mismatches:	30
Query Match:	87.63%	Indels:	0
DB:	5	Gaps:	0

US-10-031-331B-40 (1-473) x AAD16797 (1-2235)

Qy	1	HsThrValAspLeuThrileGluAlaMetMetLeuAspSerGlnAspSerAspLeuAsp	20
Db	317	CACCTCTGTTGATTTGACCTGTTGAGGCTATGATGCTTCAACAGCTTCAGATCTTGAC	376
Qy	21	LysGluGluArgProGluLeuLeuSerMetLeuProLeuGluGluLysCysLeuLeu	40
Db	377	AAAGTGGAGGACCTGAGGACTTCCATGCTCCACCTTATGAAGGAAAGTCTGCTTA	436
Qy	41	GluLeuGluValGlyLeuGlyArgPheThrGlyGluLeuAlaGluValGlyGlnVal	60
Db	437	GAACCTCGGTCCTGTTGTTGCTTTTACTGTTGTAATGGCCGAGAAAGCTAGCAGGTC	496
Qy	61	IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyIleTyr	80
Db	497	ATCGCTCTGGATTTTCATTTGAGAGTGTATAAAGGAAGATGAAAGCATTAATGGCATTAC	556

Qy	81	LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100
Db	557	AAATATGTAAGTTTATGCTGCTGATGACATCTCCAAAGTCTCAACATTTCCACNAAT	616
Qy	101	SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu	120
Db	617	TCCGTGGATATCATATTTCTCCAAATGGCTACTCATGTATCTTTCTGATGAAGGGTTGAG	676
Qy	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu	140
Db	677	CGTCTGGTTGAAGGATGTTGAAATGGTTGAGCCAGGAGGATACATTTCTTCAGAGAA	736
Qy	141	SerCysPheHisGlnSerGlyAspHisLysLysSerAsnProThrHisTyrArgGlu	160
Db	737	TCTTGTCTTTCATCAATCAGGAGATCAACAGCGAAGCAATCCACCCACTACCGTGA	796
Qy	161	ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180
Db	797	CCTAGGTTCTACACCAAGATCTTCAAGAATGCCATATGCAAGATGATCTCGGAATCTCC	856
Qy	181	TyrGluLeuSerLeuLeuSerCysLysCysIleGlyValTyrValArgAsnLysLysLeu	200
Db	857	TATGAGCTCTCCCTAATTTGGCTGCAATGTATTTGGAGCTTATGTCAAAGCAAGAGAAAT	916
Qy	201	GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln	220
Db	917	CAGAACCAAGATAAGCTGTTATGGCAGAAAGTTGATTCAGAGGATGACAAAGGGTTCCAG	976
Qy	221	ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe	240
Db	977	CGATTTCTGGATTTCTAGTCAATCAAGTTTAAACAGCATACTGCTTATGAGGCTGTTT	1036
Qy	241	GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu	260
Db	1037	GCTCTGTTGATGTATGTACCGAGGACTCGAACAACCAAGAGGTTGTATCAAGCTT	1096
Qy	261	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyLysAspPhe	280
Db	1097	GACTTGAAGCCTGGCCAGAAAGCTCCTAGATGTGGTGTGGTATAGGTGAGGTGATTTT	1156
Qy	281	TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle	300
Db	1157	TACATGGCAGAGAACTATGATGTTGGAGTTGTTGGAATTTGATCTCTCCATTATATGAT	1216
Qy	301	SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp	320
Db	1217	TCTTTTGGCTTTGAGCGCTCAATTTGGCTCAATGCTGTTGAGTTGAGGTGGCAGAT	1276
Qy	321	CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle	340
Db	1277	TGCACCAAGAAAGATTAACCTCGAAACTCTTTTATGATGTCATCTACAGCGGTGATACCAT	1336
Qy	341	LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly	360
Db	1337	CTGCATATTGAGGACAAACCTGCTTTATTAGATCTTCCAAATGTTGAAACCTGGA	1396
Qy	361	GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla	380
Db	1397	GGCAAGTCTTATTATGACTACTGTAAAGAGTGTGGTACACCTTCAGCTGAAATTTGCT	1456
Qy	381	AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu	400
Db	1457	GCATACATCAGGACAGAGGGATATGATCTCCAGATGTGAAGGCATATGGCAAGATGCTT	1516
Qy	401	LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal	420
Db	1517	AAAGATGCTGGAATTTGTTGAGGTTATGTGCTGAAGATAGGACTGACCACTTCAAGTT	1576
Qy	421	LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer	440
Db	1577	CTGCAGAGGAAGTATGATGCTTCTGAACAGGAGAGAGTACTTCATTGATGATTTCTCT	1636
Qy	441	GluGluAspTyrAsnAspIleValGlyGlyTyrPheAsnAspLysLeuArgThrAlaLys	460

Db 1637 GAGGAGGATTATACGACATAGTTGATGGTTGGAAGGCCAGTGTGGTGGAGCTACAGAG 1696
 QY 461 GYGLUGLNAAGTTPGLYLeuPheValAlaLysLys 472
 Db 1697 GGTGAGCAACATGGGGTTTGTTCATTTGCCAGAAA 1732

RESULT 3

AAH44262

ID AAH44262 standard; DNA; 1962 BP.

AC AAH44262;

XX 21-SEP-2001 (first entry)

DT Physcomitrella patens 78_ppprot1_092_e12-260-rev gene.

XX Tocopherol and carotenoid metabolism related protein; TCMP; synthesis;

XX Physcomitrella patens; moss; algae; microorganism; fungus; plant;

KW identification; genome mapping; modulation; evolutionary study;

KW cellular production; fine chemical; ds.

XX Physcomitrella patens.

OS WO2001144276-A2.

XX 21-JUN-2001.

XX 14-DEC-2000; 2000WO-EP012698.

XX 16-DEC-1999; 99US-0171121P.

XX (BADI) BASF PLANT SCI GMBH.

XX Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;

XX Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;

XX WPI, 2001-398121/42.

XX P-PSDB; AA899889.

XX Tocopherol and carotenoid metabolism related protein (TCMP), used to

XX produce fine chemicals, is isolated from mosses, algae, microorganisms,

XX fungi, plants, or their fragments.

XX Claim 6; Page 112-113; 123pp; English.

XX The present invention describes isolated tocopherol and carotenoid

XX metabolism related proteins (TCMP) (I) from mosses or algae, used as

XX microorganisms or fungi, plants, or its fragments. (I) can be used as

XX enzymes in the production of fine chemicals or in the metabolism of

XX tocopherols and carotenoids. (I) also assist in transmembrane transport.

XX The fine chemicals that can be produced include lipids, fatty acids,

XX vitamins, cofactors, enzymes, amino acids, and nucleotide bases.

XX Nucleotide sequences, proteins, vectors and host cells from the present

XX invention can be used: (a) to identify mosses related to Physcomitrella

XX patens; (b) in mapping genomes of mosses related to Physcomitrella

XX patens; (c) in the determination of functional TCMP regions; (f) and in the

XX cellular production of fine chemicals. AAH44262 to AAH44262 encode the

XX Physcomitrella patens TCMP proteins given in AA899849 to AA899889.

XX AAH44262 to AAH44262 represent nucleotide sequence used in the

XX exemplification of the present invention

XX

SQ Sequence 1962 BP; 517 A; 416 C; 518 G; 511 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,46e-189 Length: 1962

Score: 1829.50 Matches: 332

Percent Similarity: 83.76% Conservative: 65

Best Local Similarity: 70.04% Mismatch: 74

Query Watch: 72.74% Indels: 3

DB: 4 Gaps: 2

US-10-031-331B-40 (1-473) x AAH44262 (1-1962)
 QY 1 HistHrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 Db 415 CATCTGTGGAGCCTAGCGTTGAGGCAATGATGCTTGTTCAGGCGCTCCAAACTCGAT 474
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu 40
 Db 475 AAAGAGAGAACGACCCGAGATTTTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
 Db 535 GAGCTCGGAGCAGGATCGGTGCTTACTGTGCTGAGCTTGCAGAGCATGTCAGGTTCATGTG 594
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysValLeuValIleAsnGlyHisTyr 80
 Db 595 CTTCGCGATGTTTCTGAGATCTCATCAAGAAGACGAGGATGTGAACGCTCACTAC 654
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
 Db 655 AACACATCGATTTCAATGTGCGATGAGCTCTCCAGACCTGATATATTCACACAGGT 714
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120
 Db 715 TCTCGGATCTCGTGTGTTTCAATTTGCTTCTCATGTACTTCTGCTGACGAGGTTAAA 774
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
 Db 775 GGTCTAGCATCACCGTTATGGAGTGGCTCAGGCTTGGAGGATACATTTCTTCAGAAA 834
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
 Db 835 TCCTGCTTCCACCATCAGGATCAACAGCGAAGAACAACTCTACTCCTACCTCCTCAA 894
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 Db 895 CCCAACGAGTACAGCAATCTTCCAGCAGCGCTACATCGAAGAG---GATGGGTCTAT 951
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysValAsn 200
 Db 952 TTCAGTTGGAATGCTGGATGCAATGTGCTGACATCATCTGCGAATATGAGAGAAAT 1011
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysVal-----AspSerLysAspLysGly 218
 Db 1012 CAAAACCCAGGTGTGTGTTATGGAGGAAAGTTCAGTCCGATCGGATCGGAGCGGTGT 1071
 QY 219 PheGlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArg 238
 Db 1072 TTCAGAGAGTTTGGACACCCACAGTACACGTCACCTGGAATCTGCGTTACGAGGT 1131
 QY 239 ValPheGlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSer 258
 Db 1132 ATTTTGGAGAGGATTTGTTAGCAGCGGTGGATCGAAACCAACCAAGAGCTTTTGAAGT 1191
 QY 259 MetLeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGly 278
 Db 1192 ATGCTGCACTTGAAGCCAGGACAGCGTCTCTGAGCTTGGATGCTGGATCGGAGGTGT 1251
 QY 279 AspPheTyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsn 298
 Db 1252 GATTTCTACATGCGCGAGATATGATGCTGAAGTTGTTGGCATCGCTGCTCTAAAT 1311
 QY 299 MetIleSerPheAlaLeuGluArgSerIleGlyLysCysAlaValGluPheGluVal 318
 Db 1312 ATGATTTGCTTGTCTTGAACGATCGATCGGAGAAATGTGCACTCGAGTTGAAGTT 1371
 QY 319 AlaAspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAsp 338
 Db 1372 GGGATTTGACCAACAGATTAATTTACCTCAGCATCTTTTGTATCTCATCTACAGTCTGAT 1431
 QY 339 ThrIleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrValTrpLeuLys 358
 Db 1432 ACCATTTCTACATTCAGATAAACCTCGGCTTTTTCACGCGTTTATATATGTTGAGAG 1491

```

QY 359 ProGlyGlyLysValLeuLeuSerAspTyrCysLysLysAlaGlyProProSerProGlu 378
DB 1492 CTGGAGGTCGGTGCTTATCAGTACTGTAGAGCTCCCAAACTCCGTGGCGAG 1551
QY 379 PheAlaAlaTyrLleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGln 398
DB 1552 TTCGTGTCATATTCACGAGAGGGTTATGATCTCCATAGCGTTTCAGAGTACGAGAG 1611
QY 399 MetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIle 418
DB 1612 ATGCTGGGAAGATGCGGTTTGTGGAAGTGTGTCAGAGGACCGCACGATCAGTTCAAT 1671
QY 419 ArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAsp 438
DB 1672 GAAGTGTTCACAGAGGAGTACGCCACCTGAGCAGGTCGTGACCGTTCATCAACGAT 1731
QY 439 PheSerGluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgThr 458
DB 1732 TTCCTCGAGGAGGATTAATCACTATGTCAGCGGATGGAAGATGAGTAAAGCTGAAGCGCTGT 1791
QY 459 AlaLysGlyGluGlnArgTyrGlyLeuPheValAlaLysLys 472
DB 1792 TCGAATGACGAACAGAGTGGGACTCTTCATAGCCTCAAG 1833

RESULT 4
ID AAD16798 standard; DNA; 1120 BP.
XX AAD16798;
XX 29-NOV-2001 (first entry)
XX Spinach PEAMT truncated DNA.
XX Spinach; S-adenosyl-L-methionine, phosphoethanolamine N-methyltransferase;
KW PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;
KW phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;
KW glycine betaine; choline-O-sulphate; lipid content alteration;
KW osmotic stress tolerance; nutritional value; transgenic plant;
KW cryoprotectant; ds.
XX Spinacia oleracea.
XX Key Location/Qualifiers
FH CDS 254..1120
FT /*tag= a
FT /*product= "Spinach PEAMT truncated protein"
FT /*note= "CDS does not include stop codon"
FT /*partial
XX WO200168870-A2.
XX 20-SEP-2001.
XX 15-MAR-2001; 2001WO-US008352.
XX 15-MAR-2000; 2000US-00525885.
XX (UYFL) UNIV FLORIDA.
XX (UYCA-) UNIV CARNEGIE MELLON.
XX Hanson AD, Nuccio ML, Henry SA;
XX WPI; 2001-565796/63.
XX P-PSDB; AAS09761.
XX New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
XX polypeptides, useful for modulating the levels of cellular intermediates
XX such as phosphedimethylethanolamine and for altering the lipid content in
XX plants cells.
XX Claim 18; Page 109; 158pp; English.

```

XX The present sequence is spinach S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase (PEAMT) truncated DNA. The PEAMT sequences are useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine, phosphomono-methylethanolamine, choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine betaine. They are useful for altering the lipid content in plant cells. The polynucleotides are also useful for improving the osmotic stress tolerance of a plant and increasing the cryoprotectant properties of a plant. The present invention also relates to methods and compositions comprising PEAMT used for generating transgenic plants with increased nutritional value

XX SQ Sequence 1120 BP; 304 A; 218 C; 256 G; 342 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,41e-128 Length: 1120
Score: 1270.00 Matches: 235
Percent Similarity: 94.34% Conservative: 15
Best Local Similarity: 88.68% Mismatches: 15
Query Match: 50.50% Indels: 0
DB: 5 Gaps: 0

US-10-031-331B-40 (1-473) x AAD16798 (1-1120)

```

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
DB 317 CACTCTGTTGATTTGACTGTTGAGGCTATGATGCTTGATTCACAGCTTCAGATCTTGAC 376
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGluLysCysLeuLeu 40
DB 377 AAAGTGGAGCGACCTGAGGTACTTTCATGCTTCCACCTTATGAAGGAAGTCTGCTTA 436
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
DB 437 GAACCTCGGTGCTGATGATGCTGCTTTACTGGTGAATTCGCCGAGAAAGCTAGCCAGGTC 496
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
DB 497 ATCGCTCTGGATTTCAATTGAGAGTGTATATAAGAGATGAAGCATTAATGGGCATTAC 556
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPheHis 100
DB 557 ABAATGTGAAGTTATGCTGCTGATGTGCATCTCCAGTCTCAACATTCACCAAT 616
QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120
DB 617 TCCGTGGATATCATATTCCTCAATTCGCTACTCATGTATCTTTCTGATGAAGAGGTTGAG 676
QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheArgGlu 140
DB 677 CGTCTGTTCAAGGATGTTGAATGTTGAAGCCAGGAGGATACATTTCTTCAGAGAA 736
QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
DB 737 TCTTGTGTTTCATCAATCAGGAGATCAAGCGCAAGCAATCCCAACCTACCTACCGTGA 796
QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
DB 797 CCTAGGTTCTACACCAAGATCTTCAAGAATGCCATATGCAAGATGATCTTCGGGAATCC 856
QY 181 TyrGluLeuSerLeuLeuSerCysLysCysGlyIleGlyValTyrValArgAsnLysAsn 200
DB 857 TATGAGCTCTCCCTCAATTCGCTCAATGATTTGGAGCTTATGTCANAGCAAGAGAT 916
QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
DB 917 CAGAACCAAGATAGCTGTTATGCGCAGAAAGTTGATTCAGAGATGACACAGGGTTCAG 976
QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
DB 977 CGATCTCTGGATCTAGTCAATCAAGTTTACAGCATCTACTCGTTATGAGCGTGATTT 1036

```

QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 Db 1037 GGTCTGGTATGTTAGTACCGGAGGACTCGAACAACCAACGAGGTTGTATCAAGCTT 1096

QY 261 AspLeuValProGly 265
 Db 1097 GACTTGAAGCCCTGGG 1111

RESULT 5

ACL18332/c
 ID ACL18332 standard; DNA; 636 BP.

XX AC ACL18332;

XX XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #8323.

DE Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KW gene; ss.

XX OS Hordeum vulgare; ssp. spontaneum.

XX PN WO2003057877-A1.

XX XX 17-JUL-2003.

XX PF 16-DEC-2002; 2002WO-IB005403.

XX PR 20-DEC-2001; 2001JP-00387059.

XX PR 20-DEC-2001; 2001JP-00387131.

XX PR 20-DEC-2001; 2001JP-00403299.

XX PR 20-DEC-2001; 2001JP-00403300.

XX PR 27-SEP-2002; 2002JP-00327515.

XX PA (UYN1-) UNIV JAPAN OKAYAMA.

XX PI Sato K, Takeda K, Kohara Y;

XX DR WPI; 2003-587127/55.

XX PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.

XX PS Disclosure; SEQ ID XX; 284pp; Japanese.

XX CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)

XX SQ Sequence 636 BP; 154 A; 179 C; 138 G; 165 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8-81e-89	Length:	636
Score:	905.00	Matches:	162
Percent Similarity:	86.26%	Conservative:	20
Best Local Similarity:	76.78%	Mismatches:	29
Query Match:	35.38%	Indels:	0
DB:	8	Gaps:	0

US-10-031-331B-40 (1-473) x ACL18332 (1-636)

QY 125 ArgMetLeuLysTrpLeuLysProGlyGlyTyrLysPhePheArgGluSerCysPheHis 144
 Db 634 AGAATAGTGAAGTGGTGGAGCCCTGGTGCACATATTTTATCAGGAATCGTGTCCAT 575
 QY 145 GlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGluProArgPheTyr 164
 Db 574 CAGTCTGGTGAATCCAAAGAGGAAGTGAACCCAGCCACTACCCGGAGCCGAGGTTTAC 515
 QY 165 ThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSerTyrGluLeuSer 184
 Db 514 ACCAAGGTGTTCAAGGAATGCCACTCCATATGACCAAGAGGGGAACCTCTTTAGAGCTTCT 455
 QY 185 LeuLeuSerCysLysCysValLeuGlyAlaTyrValArgAsnLysLysAsnGlnLeu 204
 Db 454 CTGGTAACCTCCAGTGTATTTGGAGCTTATGTGAAAGCAGAGAGAACCCAGACGATA 395
 QY 205 SerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGlnArgPheLeuAsp 224
 Db 394 TCTCGCTATGGGAGAGGTCAGTGCACAGAGACAAAGGCTTTTCAGAGATTCCTGGAC 335
 QY 225 ThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheGlyProGlyTyr 244
 Db 334 AATGTGCAGTACAAATCCACTCGAATCTTGGCTACGAGCGGTGTTCGGGAGGTTAT 275
 QY 245 ValSerThrGlyGlyTyrGluThrLysGluPheValSerMetLeuAspLeuLysPro 264
 Db 274 GTTAGCACCGTGGATTTCGAGACCAACAAAGGATTTTGTGCAAGCTGACCTGAAGCT 215
 QY 265 GlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysPheTyrMetAlaGlu 284
 Db 214 GGCCAGAGAGGTCTTGTATGTTGGGTGGCATCGAGAGGCGCACTTCTACATGGCTGAA 155
 QY 285 ThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeu 304
 Db 154 ACCTATGATGTCATGCTCTCGCATCATCTTTCCATCAACATGCTTCTTCTTGGCATC 95
 QY 305 GluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAspCysThrLysIle 324
 Db 94 GAGCGTGCATCGGTGGTCTCTGCTCGGTGAGTTGAGTTGAGTTGCTGCTGCTGCTGCTGCT 35
 QY 325 AsnTyrProAspAsnSerPheAspValIleTyr 335
 Db 34 GAATACCCAGAGACACAGCTTCGATGTCATCTAC 2

RESULT 6

ACL25710
 ID ACL25710 standard; DNA; 605 BP.

XX AC ACL25710;

XX DT 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #15701.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KW gene; ss.

XX OS Hordeum vulgare; var. (cul.Haruna Nijo).

XX FN WO2003057877-A1.

XX PD 17-JUL-2003.

XX PF 16-DEC-2002; 2002WO-IB005403.

XX PR 20-DEC-2001; 2001JP-00387059.

XX PR 20-DEC-2001; 2001JP-00387131.

XX PR 20-DEC-2001; 2001JP-00403299.

XX PR 20-DEC-2001; 2001JP-00403300.

XX PR 27-SEP-2002; 2002JP-00327515.

Db 514 CCAAGTTTACACTAAGCTATTTAAAGAGGCGCCATGCTGATCAGAGTGCGAGCTCC 573

Qy 181 TyrGluLeuSerLeuSerCysAlaValGluPheGluValAlaAaspCysThrLys 323
 Db 713 CTGAGCATGCTATTGGAGCGCAAGTGGCGAGTTCGAGTTGAGTTGCTGATTCACACG 654

Qy 324 IleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIle 343
 Db 653 AAGACATACCCAGACAAATACATTGATGTTATCTACAGCCGTGACACCATCTTCACATA 594

Qy 344 GlnAspLysProAlaLeuPheArgSerPheTyrIleTyrTrpLeuLysProGlyGlyVal 363
 Db 593 CAAAGTAAACCCGCTTTGTTTAGAAGTTCTTCAATGGCTAAACCTGGTGAAGTTC 534

Qy 364 LeuIleSerAspTyrCysLysLysAlaGlyProSerProGluPheAlaAlaTyrIle 383
 Db 533 CTAATCAGCGATTACTGTAGGAGTCCAGGAACCATCTGACAGGTTTGGCTATACATT 474

Qy 384 LysGlnAspGlyTyrAspLeuHisValLysGluTyrGlyGlnMetLeuLysAla 403
 Db 473 AAGCAGAGGGGTACGACCTTCATGATGTAGAGACTTATGACAGATGCTAGAGAATGCT 414

Qy 404 GlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLys 423
 Db 413 GGTTCATGATGTCTATGCGAAGACCGCAGTCAGTCTCTGAAAGTTCATGACAGG 354

Qy 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443
 Db 353 GAGCTAGCGGAAGTTGAAAGAACAAAGATGATTTTCTAGCCGACCTTGGTCAGGAGG 294

Qy 444 TyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGln 463
 Db 293 TATGACGATATTGTGACTGGTGGTGAACCGCAAACTTCATAGAGCTCTGCTGGTGAGCAG 234

Qy 464 ArgTrpGlyLeuPheValAlaLysLys 472
 Db 233 AGGTGGGGCTGTTTCATTGGGACCAAG 207

RESULT 10
 ACL18347/c
 ID ACL18347 standard; DNA; 575 BP.
 XX
 AC ACL18347;
 XX
 DT 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 XX
 DE DNA clone originating in barley containing SNP encoding sequence #8338.
 XX
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KW gene; ss.
 XX
 OS Hordeum vulgare; var. (cul.Haruna Nijo).
 XX
 FN WO2003057877-A1.
 XX
 PD 17-JUL-2003.
 XX
 PF 16-DEC-2002; 2002WO-IB005403.
 XX
 PR 20-DEC-2001; 2001JP-00387059.
 PR 20-DEC-2001; 2001JP-00387131.
 PR 20-DEC-2001; 2001JP-00403299.
 PR 20-DEC-2001; 2001JP-00403300.
 XX
 XX 27-SEP-2002; 2002JP-00327515.
 PA (UTNI-) UNIV JAPAN OKAYAMA.
 XX
 PI Sato K, Takeda K, Kohara Y;
 XX
 DR WPI; 2003-587127/55.
 XX
 PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley

Disclosure; SEQ ID XX; 284pp; Japanese.

The present invention relates to oligonucleotide clones originating in barley (Hordeum vulgare) which contain single nucleotide polymorphisms (SNP). The oligonucleotides may be used for analysis of SNPs among barley varieties, identification of particular varieties and creation of new varieties by analysis, isolation of specific genes and production of new varieties by transformation of barley varieties with them and production of new barley varieties with desired properties. The present sequence represents an oligonucleotide clone DNA sequence featured in the specification. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published-pct-sequences

Sequence 777 BP; 230 A; 188 C; 146 G; 212 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	773e-74	Length:	777
Score:	769.00	Matches:	139
Percent Similarity:	86.77%	Conservative:	25
Best Local Similarity:	73.54%	Mismatches:	25
Query Match:	30.58%	Indels:	0
DB:	8	Gaps:	0

US-10-031-331b-40 (1-473) x ACL19405 (1-777)

Qy 284 GluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIleSerPheAla 303
 Db 773 GAAACATATGATGTTATGTTGTCATGATGCTTTCATTAACATGATGTTTCATTGCA 714

Qy 304 LeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAaspCysThrLys 323
 Db 713 CTGAGCATGCTATTGGAGCGCAAGTGGCGAGTTCGAGTTGAGTTGCTGATTCACACG 654

Qy 324 IleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIle 343
 Db 653 AAGACATACCCAGACAAATACATTGATGTTATCTACAGCCGTGACACCATCTTCACATA 594

Qy 344 GlnAspLysProAlaLeuPheArgSerPheTyrIleTyrTrpLeuLysProGlyGlyVal 363
 Db 593 CAAAGTAAACCCGCTTTGTTTAGAAGTTCTTCAATGGCTAAACCTGGTGAAGTTC 534

Qy 364 LeuIleSerAspTyrCysLysLysAlaGlyProSerProGluPheAlaAlaTyrIle 383
 Db 533 CTAATCAGCGATTACTGTAGGAGTCCAGGAACCATCTGACAGGTTTGGCTATACATT 474

Qy 384 LysGlnAspGlyTyrAspLeuHisValLysGluTyrGlyGlnMetLeuLysAla 403
 Db 473 AAGCAGAGGGGTACGACCTTCATGATGTAGAGACTTATGACAGATGCTAGAGAATGCT 414

Qy 404 GlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLys 423
 Db 413 GGTTCATGATGTCTATGCGAAGACCGCAGTCAGTCTCTGAAAGTTCATGACAGG 354

Qy 424 GluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSerGluGluAsp 443
 Db 353 GAGCTAGCGGAAGTTGAAAGAACAAAGATGATTTTCTAGCCGACCTTGGTCAGGAGG 294

Qy 444 TyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgArgThrAlaLysGlyGluGln 463
 Db 293 TATGACGATATTGTGACTGGTGGTGAACCGCAAACTTCATAGAGCTCTGCTGGTGAGCAG 234

Qy 464 ArgTrpGlyLeuPheValAlaLysLys 472
 Db 233 AGGTGGGGCTGTTTCATTGGGACCAAG 207

RESULT 10
 ACL18347/c
 ID ACL18347 standard; DNA; 575 BP.
 XX
 AC ACL18347;
 XX
 DT 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 XX
 DE DNA clone originating in barley containing SNP encoding sequence #8338.
 XX
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KW gene; ss.
 XX
 OS Hordeum vulgare; var. (cul.Haruna Nijo).
 XX
 FN WO2003057877-A1.
 XX
 PD 17-JUL-2003.
 XX
 PF 16-DEC-2002; 2002WO-IB005403.
 XX
 PR 20-DEC-2001; 2001JP-00387059.
 PR 20-DEC-2001; 2001JP-00387131.
 PR 20-DEC-2001; 2001JP-00403299.
 PR 20-DEC-2001; 2001JP-00403300.
 XX
 XX 27-SEP-2002; 2002JP-00327515.
 PA (UTNI-) UNIV JAPAN OKAYAMA.
 XX
 PI Sato K, Takeda K, Kohara Y;
 XX
 DR WPI; 2003-587127/55.
 XX
 PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley

QY 222 cLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPheGlyPr 242
 Db 384 TTGGATTAATGTCAGTACAAACCAAGTGAATATTACCTATGAGCTGTTTGGCA 325
 QY 242 cGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeuAspLe 262
 Db 324 AGTTTGTGAGCAGCTGTTGAATGAGACTACAAAGAAATTTGGACTTCTGGATCT 265
 QY 262 wLysProGlyGlnPheValLeuAspValGlyCysGlyIleGlyGlyGlyAspPheTyrMe 282
 Db 264 TAAACCTGGCAGAGTCTGATGTTGATGTAATCGGGGTGGATTTTAT 205
 QY 282 tAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIleSerPh 302
 Db 204 GCGCGAAACTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 145
 QY 302 cAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAspCysTh 322
 Db 144 TGCATCGAGCATCTATTGGACGCAAGTGGCGATGCGAGTTTGATGTTGATGTTGAT 85
 QY 322 rLysIleAsnTyrProPheAsnSerPheAspValIleTyrSerArgAspThrIleLeuHi 342
 Db 84 CAAGAGACATACCCAGACATACATTTGATGTTATCTACAGCGGTGACACCATCTTCA 25
 QY 342 sIleGlnAspLysProAlaLeu 349
 Db 24 CATACAGATAAACCGCTTGG 3
 RESULT 12
 ACL18342/c
 ID ACL18342 standard; DNA; 650 BP.
 XX
 AC ACL18342;
 XX
 17-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 XX
 DE DNA clone originating in barley containing SNP encoding sequence #8333.
 XX
 KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 XX
 OS Hordeum vulgare; var. (cul.Haruna Nijo).
 XX
 FN WO2003057877-A1.
 XX
 PD 17-JUL-2003.
 XX
 PF 16-DEC-2002; 2002WO-1B005403.
 XX
 PR 20-DEC-2001; 2001JP-00387059.
 PR 20-DEC-2001; 2001JP-00387131.
 PR 20-DEC-2001; 2001JP-00403299.
 PR 20-DEC-2001; 2001JP-00403300.
 PR 27-SEP-2002; 2002JP-00327515.
 XX
 PA (UYN1-) UNIV JAPAN OKAYAMA.
 XX
 PI Sato K, Takeda K, Kohara Y;
 XX
 DR WPI; 2003-587127/55.
 XX
 PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.
 XX
 PS Disclosure; SEQ ID XX; 284pp; Japanese.
 XX
 CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype

CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 650 BP; 123 A; 194 C; 177 G; 156 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,31e-71 Length: 650
 Score: 747.50 Matches: 137
 Percent Similarity: 86.17% Conservative: 25
 Best Local Similarity: 72.87% Mismatches: 25
 Query Match: 29.72% Indels: 1
 DB: 8 Gaps: 1

US-10-031-331B-40 (1-473) x ACL18342 (1-650)

QY 1 KieThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 Db 564 CACTCCAGAGACCTCACCGTCGAGTCGATGCTCGCTCCGCGCCAGACCTCGAC 505
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40
 Db 504 AAGGAGGAGCGCCGAGAGGTGCTGGCCATCTCTATGAGGGCAAGACTGTGCTG 445
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
 Db 444 GAGCTCGCGCGCATCGCGCTTTACCGCGAGCTGGCCAGAGGCGCGCCACGTC 385
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHis 79
 Db 384 ATCCGCTGAGCTTCATCGACGCTCATCAGAGAAACGAGAGATCAATGGGCAATC 325
 QY 80 TyrLysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPro 99
 Db 324 CACAGCAACATACCTTCATGTCGCGCGAGCTGAGCTGCGCGAGCTCAAGATCGAGGAC 265
 QY 100 HisSerLeuAspValIlePheSerAsnTyrLeuLeuMetTyrLeuSerAspGluVal 119
 Db 264 AACTCCATCGACATCGCTTCCTCAACTGCTGCTCATGTACTCATCAAGAGAGGT 205
 QY 120 GluAsnLeuValGluArgMetLeuLysTrpLeuLysProGlyTyrIlePhePheArg 139
 Db 204 GAGAAGCTGATTGGCAGATAGTGAAGTGGCTGAAGCCTGGTGGACATATTTTATCAG 145
 QY 140 GluSerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArg 159
 Db 144 GAATCGTGTCTCCATCAGTCTGCTGATTCCAGAGAGAAAGTGAACCCGACCTACCG 85
 QY 160 GluProArgPheTyrThrIysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsn 179
 Db 84 GAGCGAGGTTTACACCAAGGTTTCAAGGAATGCCACTCTCTATGACCAAGAGGGGAAAC 25
 QY 180 SerTyrGluLeuSerLeuLeuSer 187
 Db 24 TCTTTTGAGCTTCTCTGTTAACT 1

RESULT 13

ACL18325/c

ID ACL18325 standard; DNA; 622 BP.

XX

AC ACL18325;

XX

27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX

DE DNA clone originating in barley containing SNP encoding sequence #8316.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KW

KW gene; ss.
 XX Hordeum vulgare; esp. spontaneum.
 XX WO2003057877-A1.
 PN 17-JUL-2003.
 PD 16-DEC-2002; 2002WO-IB005403.
 PP 20-DEC-2001; 2001JP-00387059.
 PR 20-DEC-2001; 2001JP-00387131.
 PR 20-DEC-2001; 2001JP-00403299.
 PR 20-DEC-2001; 2001JP-00403300.
 PR 27-SEP-2002; 2002JP-00327515.
 XX (UYNI-) UNIV JAPAN OKAYAMA.
 XX Sato K, Takeda K, Kohara Y;
 PI WPI; 2003-587127/55.
 DR
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.
 XX
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 XX
 CC The present invention relates to oligonucleotide clones originating in
 CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ip.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 622 BP; 118 A; 187 C; 173 G; 144 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 6.02e-70 Length: 622
 Score: 732.00 Matches: 135
 Percent Similarity: 87.10% Conservative: 27
 Best Local Similarity: 72.58% Mismatches: 24
 Query Match: 29.11% Indels: 1
 DB: 8 Gaps: 0
 US-10-031-331b-40 (1-473) x ACL18325 (1-622)
 QY 287 AspValGluValValGlyPheAspLeuSerValAsnMetIleSerPheAlaLeuGluArg 306
 DB 619 GATGTCATGCTGCTGGCATGATCTTTCTC-AACATGTTTCTTCTCGCATCGAGGT 561
 QY 307 SerIleGlyLeuLysCysAlaValGluPheGluValAlaAspCysThrLysIleAsnTyr 326
 DB 560 GCCATCGTTCGCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 501
 QY 327 ProAspAsnSerPheAspValIleTyrSerArgAspThrIleLeuHisIleGlnAspLys 346
 DB 500 CCAGAGMACATGCTGATGTCTATCATCAGCGCGTGACACCATCTCCATCATTCAGCAGAG 441
 QY 347 ProAlaLeuPheArgSerPheTyrIleTyrPheLysProGlyLysValLeuIleSer 366
 DB 440 CTTGCTCTGTTTCAGAACTTCTTCAGTGGCTGAGGCTGGTGGTGGTGGTGGTGGTGGTGGT 381
 QY 367 AspTyrCysLysIleAlaGlyProPheSerProGluPheAlaIleTyrIleLysGlnArg 386
 DB 380 GACTACTGACGAGGCGCTGGGACGCCATCAGAGGATTCCTGCTGATCATCAGCAGAGA 321

QY 387 GlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeuLysAspAlaGlyPheVal 406
 DB 320 GGCTATGACCTCCATGATGTGAAGACCTACGGAAGATGCTTGAGGATGCGGGTTCCAC 261
 QY 407 AspValLeuAlaGluAspArgThrGluGlnPheIleArgValLeuArgLysGluLeuGlu 426
 DB 260 GACGTTCATCGCGGAGACCCGACCGACCACTTCTTCTGAGGGTCTTGGAGGAGGAGCTGGGC 201
 QY 427 ThrValGluLysGluLysAspValPheIleSerAspPheSerGluLysAspTyrAsnAsp 446
 DB 200 GAGACCGAGAGAACAGAGGAGGCTTCTTGGCGACTTTCAGCCAGGAGGACTACGAC 141
 QY 447 IleValGlyGlyTrpAsnAspLysLeuArgArgThrAlaLysGlyGluGlnArgTyrGly 466
 DB 140 ATCGTCAACGGGTGGAGCGCGAGCTGAAGCGGAGCTCCGCGCGGAGCAGAGTGGGG 81
 QY 467 LeuPheValAlaLysLys 472
 DB 80 CTGTTTCATCGGACCAAG 63
 RESULT 14
 ADEB2005
 ID ADEB2005 standard; cDNA; 487 BP.
 XX
 AC ADEB2005;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Arabidopsis thaliana expressed polynucleotide seq id 776.
 XX
 KW genetically modified organism; transgenic organism; plant;
 KW inhibitor testing; activator testing; modifier testing; fungicide;
 KW insecticide; genetic function; genetic regulation; cellular metabolism;
 KW gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2003115639-A1.
 XX
 PD 19-JUN-2003.
 XX
 PF 26-JAN-2001; 2001US-00770961.
 XX
 PR 27-JAN-2000; 2000US-0178466P.
 XX
 PA (GORL/) GORLACH J.
 PA (ANY/) AN Y.
 PA (HAMI/) HAMILTON C M.
 PA (PRIC/) PRICE J L.
 PA (RAIN/) RAINES T M.
 PA (YUY/) YU Y.
 PA (RAME/) RAMEKA J G.
 PA (PAGE/) PAGE A.
 PA (MATH/) MATHW A V.
 PA (LEDF/) LEDFORD B L.
 PA (WOES/) WOESSNER J P.
 PA (HAAS/) HAAS W D.
 PA (GARC/) GARCIA C A.
 PA (KRIC/) KRICKER M.
 PA (SLAT/) SLATER T.
 PA (DAVI/) DAVIS K R.
 PA (ALLE/) ALLEN K.
 PA (HOFF/) HOFFMAN N.
 PA (HURB/) HURBAN P.
 XX
 PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
 PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
 PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
 PI Hurban P;
 XX
 WPI; 2003-810930/76.
 DR
 XX

PT Novel Arabidopsis thaliana nucleic acids useful for generating
PT genetically modified transgenic organisms, for screening biologically
XX active agents such as fungicides, insecticides.

PS Claim 1: SEQ ID NO 776; 44pp; English.

XX The invention describes a nucleic acid (I) comprising a sequence capable
CC of hybridising under stringent conditions to any one of 999 fully defined
CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,
CC 1137, 455, 219, 472 nucleotides, etc. or its fragment. (I) is useful as a
CC hybridisation probe to complementary molecules in a cDNA library. (I) is
CC also useful for generating genetically modified and transgenic organisms,
CC usually plant cells and plants. A protein encoded by (I) is useful in
CC screening assays to determine the effect of candidate inhibitors,
CC activators or modifiers of the gene product. The protein is also useful
CC for screening biologically active agents e.g., fungicides and
CC insecticides. A genetically modified cell, comprising an exogenous
CC nucleic acid, where the nucleic acid comprises transcription regulatory
CC sequences operably linked to a sequence capable of hybridising under
CC stringent conditions to (I) is useful in the study of genetic function
CC and regulation, for alteration of the cellular metabolism and for
CC screening compounds that may affect the biological function of the gene
CC or gene product. This sequence represents an Arabidopsis thaliana
CC polynucleotide of the invention.

XX SQ Sequence 487 BP; 129 A; 88 C; 131 G; 139 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,1e-68 Length: 487
Score: 719.00 Matches: 133
Percent Similarity: 89.51% Conservative: 12
Best Local Similarity: 82.10% Mismatches: 17
Query Match: 28.59% Indels: 0
DB: 9 Gaps: 0

US-10-031-331B-40 (1-473) x ADS82005 (1-487)

QY 182 GluLeuSerLeuLeuSerCysLeuValGlyAlaValValArgAnLysLysAenGln 201
DB 2 GAGCTCTCTATGTTGGTGGTCAATGCAATGGGGCTTATGTGAGAACCAAGAATCAG 61
QY 202 AenGlnLysSerLeuThrGlnLysValLeuValSerLysAspLysGlyPheGlnArg 221
DB 62 AATCAGATTTCCTGGATATGCAAAAAGTCAGCGTCGAGATGACAGGATTTCCAGCGT 121
QY 222 PheLeuAspThrSerGlnLysCysAenSerLysLeuArgGlyGluArgValPheGly 241
DB 122 TCTTGGACATGTTCAATACAGCTAGTGGGATCTTGGCTATGAGCGTGTCTTTGGG 181
QY 242 ProGlyThrValSerThrGlyGlyCysGluThrLysGluPheValSerMetLeuAsp 261
DB 182 GAGGATATGTGACACTGGTGGATTTGAGACAACTAAAGATTTGTGGCAAGATGGAC 241
QY 262 LeuLysProGlyGlnLysValLeuAspValGlyCysGlyLysGlyGlyAspPheThr 281
DB 242 CTTAAACCGGACAGAAAGCTCTAGATGTTGGTGTGATCGTGGAGGTGACTTCTAC 301
QY 282 MetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValMetLysSer 301
DB 302 ATGCTTGAAATTTGAGATGTCATGTTTGGATTCGATCTGTGGTCAATGATCTCT 361
QY 302 PheAlaLeuGluArgSerLysGlyLeuLysCysAlaValGluPheGluValAlaAspCys 321
DB 362 TTGGCACTGGAGCGGGCCATTTGGACTCAAAATGCTCAGTGGAGTTGAAGTGGTATTC 421
QY 322 ThrLysLysLeuThrProAspAenSerPheAspValLysSerArgAspThrLeu 341
DB 422 ACCACCAAAACATATCCCGATATTTCTTGTGATGTCATTTACGCCGTGACACTTCTG 481
QY 342 HisLeu 343
DB 482 CACATC 487

RESULT 15

ACL25708
ID ACL25708 standard; DNA; 589 BP.

XX ACL25708;

XX 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

XX DNA clone originating in barley containing SNP encoding sequence #15699.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
XX gene, ss.

XX Hordeum vulgare; var. (cul.Haruna Nijo).

XX WO2003057877-A1.

XX 17-JUL-2003.

XX 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

XX 20-DEC-2001; 2001JP-00387131.

XX 20-DEC-2001; 2001JP-00403299.

XX 20-DEC-2001; 2001JP-00403300.

XX 27-SEP-2002; 2002JP-00327515.

XX (UYMI-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA

XX sequences containing them for analysis and identification of barley

XX varieties and production of barley transformants with desired

XX characteristics.

XX Disclosure; SEQ ID XX; 284pp; Japanese.

XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare), which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX SQ Sequence 589 BP; 161 A; 120 C; 166 G; 142 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.62e-66 Length: 589
Score: 697.00 Matches: 131
Percent Similarity: 87.95% Conservative: 15
Best Local Similarity: 78.92% Mismatches: 20
Query Match: 27.71% Indels: 0
DB: 8 Gaps: 0

US-10-031-331B-40 (1-473) x ACL25708 (1-589)

QY 1 HisThrValAspLeuThrLysGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20

DB 89 CACTCCAGGACCTCACTGTGTGAGGCATGAGTCTCGACTTCCTGCGCGGATCTGGAC 148

QY 21 LysGluGluArgProGluLysLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40

DB 149 AAGGAGGAGCGCCCGAGAGATGCTGCTTTTCTTCTTCTATGAGGAAATCAGTCTGT 208

Qy 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLeuAlaGlyGlnVal 60
Db 209 GAGCTTGGCGCGGAATAGTCTTACTGGAGAACTGGCTAAGACAGCTGGGCGATGTT 268
Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLeuLeuValIleAsnGlyHisTyr 80
Db 269 CTTCATGGATTTTCATGGAGTGTGATTAAGAGATGAAGCATTAATGGCCATTAC 328
Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrIleuSerPheProHis 100
Db 329 GAAATGCACTCTTCATGTGTGCTGATGTATCATCTCCAGACCTGGTGATTTGAGGACAA 388
Qy 101 SerLeuAspValIlePheSerAsnTyrLeuLeuMetTyrLeuSerAspGluValGlu 120
Db 389 TCGATTGATCTCATATTTTCAAACTGGCTACTGATGATCTTTCAGACGAGGAGTCGAG 448
Qy 121 AsnLeuValGluArgMetLeuValTyrLeuLysProGlyGlyTyrIlePheArgGlu 140
Db 449 AAGCTTGTAGAAAGAAATGTTAAATGGCTCAAGTTGGTGGCCCATCTCTTTAGAGAA 508
Qy 141 SerCysPheHisGlnSerGlyAspHisIleArgLysSerAsnProThrHisTyrArgGlu 160
Db 509 TCATGCTTCCATCATCTGGAGACTCAAAAGGAAAGTGAATCCGACACATTATCGTGAA 568
Qy 161 ProArgPheTyrThrLys 166
Db 569 CCAGGTTTTACACTAAG 586

Search completed: August 2, 2004, 00:48:18
Job time : 552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2004, 23:12:39 ; Search time 5580 Seconds
(without alignments)
3674.060 Million cell updates/sec

Title: US-10-031-331b-40
Perfect score: 2515
Sequence: 1 HTVDLTIEAWMLDSQSLD.....KLRRKANGQRMGLFVAKKK 473

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spoel/US10031331/runat 26072004 120351 14978/app query.fasta_1.647
-DB=GenBml -QFMT=fastcap -SUPPLX=ice -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODS=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us10031331@cgn 1_1_3508_orunat 26072004 120351 14978 -NCFU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

GenBml:

1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_oth:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rnd:*
36: em_hg_nam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_oth:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2515	100.0	1602	6	BD084055
2	2515	100.0	1602	6	BD093370
3	2515	100.0	1779	8	AB080186
4	2204	87.6	2235	6	AX250551
5	2204	87.6	2252	8	AF237633
6	2077	82.6	1459	8	AF133811
7	2077	82.6	1656	8	AF428454
8	2077	82.6	1660	8	AF063866
9	2063	82.0	1794	8	AF328858
10	2044.5	81.3	1870	8	AB090883
11	2018	80.2	1834	8	AF136372
12	2011	80.0	1476	8	AY091683
13	2011	80.0	1770	8	AY058175
14	2011	80.0	1809	8	AF367299
15	2011	80.0	1878	8	AF197940
16	2009	79.9	1750	8	AF193479
17	2004	79.7	1911	8	AY093093
18	1964	78.1	2069	8	AK069137
19	1918.5	76.3	1782	8	AY065971
20	1680.5	66.8	29508	8	AC079676
21	1678	66.7	95295	8	AC020889
22	1561.5	62.1	74968	8	AB019230
23	1558	61.9	3221	8	AK102037
24	1481	58.9	141715	8	AC130725
25	1481	58.9	158911	2	AC136492
26	1424	56.6	166886	8	AP003248
27	1270	50.5	1120	6	AX250553
28	1133.5	45.1	11754	2	AC138199
29	1109	44.1	105937	8	AC012679
30	979	38.9	144305	2	AC137078
31	888.5	35.3	724	8	FVU234432
32	405	16.1	286	6	AR252133
33	389	15.5	1220	3	AY429590
34	389	15.5	110000	2	PFMAL13_16
35	363	14.4	280	6	AR251334
36	347	13.8	274	6	AR248084
37	330	13.1	260	6	AR245517
38	329	13.1	31226	3	U39998
39	296	11.8	349498	1	AP003002
40	286.5	11.5	23614	3	U64834
41	286.5	11.5	140702	2	AC006888
42	262	10.4	174	6	AR250330
43	257.5	10.2	274	6	AR245829
44	241.5	9.6	846	6	AX598613
45	241.5	9.6	52101	6	AX598593

ALIGNMENTS

RESULT 1

BD084055	BD084055	1602 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD084055	1602 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Environmental stress-tolerant gene.				
ACCESSION	BD084055				
VERSION	BD084055.1	GI:22629665			
KEYWORDS	JP 2001333784-A/20.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1. (bases 1 to 1602)				
AUTHORS	Yanada, A., Ozeki, Y. and Saito, T.				
TITLE	Environmental stress-tolerant gene				
JOURNAL	Patent: JP 2001333784-A 20 04-DEC-2001;				
COMMENT	JAPAN SCIENCE AND TECHNOLOGY CORP Sueada japonica JP 2001333784-A/20 PD 04-DEC-2001 PI 19-JUL-2000 JP 2000219649 PF AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO PC C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02// PC C12N1/21, PC C12P21/08, C12N15/00, C12N5/00 CC Environmental stress-tolerant gene EH Key FT CDS Location/Qualifiers 1..1602 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"				
FEATURES	source				
ORIGIN					
Alignment Scores:	8 476-225	Length:	1602		
Pred. No.:	2515.00	Matches:	473		
Score:	100.00%	Conservative:	0		
Percent Similarity:	100.00%	Mismatches:	0		
Best Local Similarity:	100.00%	Indels:	0		
Query Match:	100.00%	Gaps:	0		
DB:	6				
US-10-031-331B-40 (1-473) x BD084055 (1-1602)					
QY	1	HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20		
Db	1	CAACCGTTGATTTAAACCAATCAAGCTATGATGCTCGATCTCAAGCTCTCGATCTTGAC	60		
QY	21	LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu	40		
Db	61	AAAGAGAACGCTCCTGAGATTCTTTCAATGCTTCGCTCTTGAAGGAAATCCCTCTTG	120		
QY	41	GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal	60		
Db	121	GAACCTGGGCTGGTATGCTGCTTTACTGTGTGAATTCGCTGAGAAAGCTGGCCAGGTT	180		
QY	61	IleAlaLeuAspPheIleGluSerAlaIleValLysLeuGlnValIleAsnGlyHisLeu	80		
Db	181	ATTGCTCTGATTTCAATGAGTGCTATCAAGAGATGAAGTATCAATGGGCACTAC	240		
QY	81	LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100		
Db	241	AAAAATGTCAGTTTATGTGCTGATGTGACTTCTCCCACTCTCAGTTTCCCAACCAT	300		
QY	101	SerLeuAspValIlePheSerAsnTrpLeuMetTrpLeuSerAspGluGluValGlu	120		
Db	301	TCATTGGATGATATCTCCCAATGTTACTCATGATATCTTCATGATGAAGAGGTGGA	360		
QY	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTrpIlePhePheAspGlu	140		
Db	361	AAATTTGGTGAAGAAATTTGAATGTTGAAGCCAGGGGTATCTTTCTTCAGAGAA	420		
QY	141	SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisLeuArgGlu	160		
Db	421	TCTTGTTCCTCATCTGGGATCAACAGCCAAAGCAATCCCAACCCCTACCTGCGTAA	480		
QY	161	ProArgPheTrpThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180		
Db	481	CCTAGGTTCTACACTAAGCCCTTCAAGAGTGTCAATTTGCAAGATGGATCTGGAACCTCT	540		
QY	181	TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTrpValArgAsnLysLysAsn	200		
Db	541	TATGAGCTCTCCCTACTTAGCTGCAAAATGATTTGGAGCTTATGTTCAGAAACCAAGAAAC	600		
QY	201	GlnAsnGlnIleSerTrpLeuTrpGluLysValAspSerLysAspAspLysGlyPheGln	220		
Db	601	CAGAACCAAGATTAGTTGGTTGGTGGCAAAAGTTGATTTTAGGATGATAGGGGTTCCAG	660		
QY	221	ArgPheLeuAspThrSerGlnTrpLysCysAsnSerIleLeuArgTrpGluArgValPhe	240		
Db	661	CGATTTCTGGATACTAGCCAGTACAAAGTGAATAGCATTTCTGCGATATAGCGGTGATTT	720		
QY	241	GlyPheGlyTrpValSerThrGlyGlyTrpGluThrThrLysGluPheValSerMetLeu	260		
Db	721	GGCCTCTGTTTATGTTAGCTCTGAGGATGATGAACCAACCAAGAGTTTGTGTCAATGCTG	780		
QY	261	AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe	280		
Db	781	GACTTGAAGCCCTGGCCAGAAAGGCTCTGGATTTGGTTGGATTTGGTGGAGGTGACTTT	840		
QY	281	TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle	300		
Db	841	TACATGGCGGAGACCTTTGATGTTGGGTTTGGATTTGATCTCTCCGTTTAATATGATT	900		
QY	301	SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp	320		
Db	901	TCCTTTGCCCTTGAGCGTTCTATTTGGCTTAATGTTGCTGTTGAGTTTGAGGTAGCAGAT	960		
QY	321	CysThrLysIleAsnTrpProAspAsnSerPheAspValIleTrpSerArgPheThrIle	340		
Db	961	TGCACCAAGATAAACTACCTCTGATTAACCTTTTGTGATGTCATCTATAGCCGTGACACAT	1020		
QY	341	LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTrpLysTrpLeuLysProGly	360		
Db	1021	CTGCATATTCAGACCAAGCCCTGCTGTTTGTAGATCTCTTACAAATGGTTGAAGCCAGGA	1080		
QY	361	GlyLysValLeuIleSerAspTrpCysLysLysAlaGlyProProSerProGluPheAla	380		
Db	1081	GGTAAAGTTCTAATCAGTGATTTACTGCAAGAAAGCTGGTCCACCTCACTGATTCGCC	1140		
QY	381	AlaTrpIleLysGlnArgGlyTrpAspLeuHisAspValLysGluTrpGlyGlnMetLeu	400		
Db	1141	GCTTACATTAAGCAGAGGGGATGATGATCTCCATGATGTAAGGAATATGGGCGAGTGC	1200		
QY	401	LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal	420		
Db	1201	AAAGATGCTGGATTTGTTGATGTTCTTCCCGAGGATAGACTGAGCGAGTTCAATCGAGTT	1260		
QY	421	LeuArgGlyGluLeuGluThrValGluLysAspValPheIleSerAspPheSer	440		
Db	1261	CTACGGAAGAACTAGAGACTGTTGAGAAGGAAGGATGTGTTCATTAGTATTTCTCT	1320		
QY	441	GluGluAspTrpAsnAspIleValGlyGlyTrpAsnAspLysLeuArgArgThrAlaLys	460		
Db	1321	GAGGAGCATTTACATGACATTTGTTGGAGTTTGGATGATAGTTGCGAGGACTGCGCAG	1380		
QY	461	GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473			
Db	1381	GGTGAGCAACGATGGGGTCTGTTCTGTTCCCAAGGAAGAG 1419			
RESULT 2					
BD093370	BD093370	1602 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093370	1602 bp	DNA	linear	PAT 27-AUG-2002
DEFINITION	Environmental stress-tolerant gene.				
ACCESSION	BD093370				
VERSION	BD093370.1	GI:22638958			
KEYWORDS	WO 0160006-A/20.				
SOURCE	unidentified				

ORGANISM unidentified
 unclassified.
 1 (bases 1 to 1602)
 YAMADA, A., Ozeki, Y. and Saito, T.
 Environmental stress-tolerant gene
 Patent: WO 0106006-A 20 25-JAN-2001;
 JAPAN SCIENCE AND TECHNOLOGY CORP., AKIYO YAMADA, YOSHIHIRO OZEKI,
 TAKEO SAITO
 OS Sueda japonica
 FN WO 0106006-A/20
 PD 25-JAN-2001
 PF 19-JUL-1999 JP 99P 235910, 24-MAR-2000 JP OOP 085377 PI
 AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO
 PC C1201/68, C12N15/29, C07K14/415, C07K16/16, C12P21/02, A01H5/00 CC
 PH Key
 PT CDS
 Location/Qualifiers
 1..1602
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Alignment Scores:
 Pred. No.: 8,47e-225 Length: 1602
 Score: 2515.00 Matches: 473
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-031-331B-40 (1-473) x BD093370 (1-1602)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 Db 1 CACACCGTGTGATTAACCAATGAGCTATGATGCTCGATTCATCAAGCTTCGATCTTGAC 60
 QY 21 LysGluGluArgProGluIleLeuSerMetLeuProProLeuGluGlyLysCysLeuLeu 40
 Db 61 AAGAAGAACGCTCTGAGATCTCTTCAATGCTTCGCGCTCTTTGAGGAAATGCTCTTG 120
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
 Db 121 GAACCTGGGGCTGGTATGCTGCTTTACTGGTGAATTCGCTGAGAAAGCTGGCCAGGTT 180
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
 Db 181 ATTGCTCTGGATTCATTTGAGAGTCTATCAGAGAGATGAAGTAAATCAATGGGCACTAC 240
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
 Db 241 AAAAATGTCAAGTTTATGTGTGCTGATGTGACTTCTCCACTCTCCAGTTTCCACACAT 300
 QY 101 SerLeuAspValIlePheSerAsnThrLeuMetTyrLeuSerAspGluGluValGlu 120
 Db 301 TCAITGGATGTGATATTCCTCAATGCTTACTCAATGATCTCTCTGATGAAGAGTGGAA 360
 QY 121 AsnLeuValGluArgMetLeuLysThrLeuLysProGlyGlyTyrIlePhePheArgGlu 140
 Db 361 AATTGGTGTGAAGAATGTGAATATGGTTGAAGCCAGAGGGGTACATTTCTTCAGAGAA 420
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
 Db 421 TCTGTGTTCCATCAATCTGGGGATCACAAAGCCAAAGCAATCCACCCACTACCGTGA 480
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 Db 481 CTTAGGTTCTACCTAAGGCTTCAAGAGTGTCAATTTGCAAGTGGATCTGGAACTCT 540
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200
 Db 541 TATGAGCTCTCCTACTAGTCTCAATATGATTGGAGCTTATGTCAGAAACACAGAAAAC 600

QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
 Db 601 CAGAACCCAGATTAGTTGGTTGGCAAAAAGTTGATTTCTTAAGCATGATAAGGGTTCAG 660
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
 Db 661 CGAATTTCTGGATACTAGCCAGTACAAAGTGTATAGCATTTCTGGATATGAGGGTGTATT 720
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 Db 721 GGCCCTGGTATGTTAGCATCTGGAGGATAGAACCCCAAGAGTGTGTGTCAATGCTG 780
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyLysAspPhe 280
 Db 781 GACTTGAAGCCCTGGCCAGAAAGTCTGGATGTGTGGTGTGTGGAAITGTGTGAGGTGACTTT 840
 QY 281 TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300
 Db 841 TACATGGCCGAGACCTTTGATGTTGAGTGTGTGGATTTGATCTCTCGTTAATATGATT 900
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 Db 901 TCCTTTGGCCCTTGAGCGCTCTATTGGGCTTAAATGGCTGTGTGAGTTTGAGGTAGCAGAT 960
 QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340
 Db 961 TGCACCAAGATAAACCTACCTGATAACTCTTTTGTGATGTCATCTATAGCCGTGACACCAAT 1020
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTyrLeuLysProGly 360
 Db 1021 CTGCATATTGAGCAAGCTGGGTTGTGTAGATCTCTTCAAAATGTTGAGGCTGAGCCAGA 1080
 QY 361 GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
 Db 1081 GGTAAAGTCTTAATAGTGAATGATTCGCAAGAAAGCTGGTCCACCTCACCTGAATTCGCC 1140
 QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
 Db 1141 GCTTACATTAAAGCAGAGGGGATGATGATCCATGATGTAAAGGATATGGCAGATGCTT 1200
 QY 401 LysAlaAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
 Db 1201 AAAGTCTGGATTTGTGTGATGTTCTTCCGAGGATAGAACTGAGCAGATTCATTCGAGTT 1260
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
 Db 1261 CTACGGAAGGAACTAGAGACTGTTCAGAGAGGAAAGGATGTGTTTCATTAGTATTCTCT 1320
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrPheAsnAspLysLeuArgGluThrAlaLys 460
 Db 1321 GAGGAGGATTACATGACATGTTGGAGTTGGATGATTAAGTTGGAGACTGCCAG 1380
 QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473
 Db 1381 GTGAGCAGCATGGGGTCTGTTCGTTCCAGAGAGAG 1419

RESULT 3
 AB080186
 LOCUS
 DEFINITION Sueda japonica PRAMT mRNA linear PLN 20-FEB-2003
 N-methyltransferase, complete cds.
 ACCESSION AB080186
 VERSION AB080186.1 GI:28436073
 KEYWORDS
 SOURCE Sueda japonica
 ORGANISM Sueda japonica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Anaradiaceae; Sueda.
 REFERENCE 1
 AUTHORS Yamada, A., Nozawa, G.T., Tanimoto, S. and Ozeki, Y.
 TITLE Glycinebetaine synthesis in Sueda japonica

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 1779)
Yamada, A., Nozawa, G.T. and Ozeki, Y.
Direct Submission
Submitted (19-FEB-2002) Akiyo Yamada, Tokyo University of
Agriculture and Technology, Department of Biotechnology; Naka-cho
2-24-16, Koganei, Tokyo 184-8598, Japan
(8-mail: yamaden@cc.tuat.ac.jp, Tel: 81-423-88-7239,
Fax: 81-423-88-7239)

FEATURES

gene	CDS
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

ORIGIN

Alignment Scores:	
Fred. No.:	9,726-225
Score:	2515.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	8
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
Matches:	473
Length:	1779

U9-10-031-331B-40 (1-473) x AB080186 (1-1779)

Qy	1	HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20
Db	178	CAACCGGTGATTTTAACCATTTGAAGCTATGATGCTCGATTCTCAAGCTTCTGATCTTGAC	237
Qy	21	LysGluGluUArgProGluIleLeuSerMetLeuProLeuGluGluLysCysLeuLeu	40
Db	238	AAAGAAGAAGCTCTGAGATTCTTTCAATGCTTCGCGCTCTTGAAGGAAAAATGCCTCTTG	297
Qy	41	GluLeuGlyAlaGlyIleGluValArgPheThrGluGluLeuAlaGluLysValaGlyGlnVal	60
Db	298	GAACTTGGGGCTGGATTATGGTCGTTTACTTGGTGAATTGGCTGAGAAAGCTGGCCAGGTT	357
Qy	61	IleAlaLeuAspPheIleGluSerAlaIleLysValysAsnGluValIleAsnGlyHisTyr	80
Db	358	ATTGCTCTCGATTTTCATTGAGAGTGTCTATCAAGAAGAATGAAGTAATCAATGGGGCACTAC	417
Qy	81	LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100
Db	418	AAAAATGTCAGAGTTTATGTGTGCTCATGTCACTTCTCCACTCTTCAGTTTCCCAACCAT	477
Qy	101	SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu	120
Db	478	TCATTGGAGTGTGATATCTCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGGAA	537
Qy	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu	140
Db	538	AAATTGGTTGAAGAATGTGTAAATGGTTGTGAAGCCAGGGGTACATTTTCTTAGAGA	597
Qy	141	SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu	160

KEYWORDS

SOURCE Spinacia oleracea (spinach)
ORGANISM Spinacia oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE Caryophyllales; Amaranthaceae; Spinacia.

Hanson, A.D., Nuccio, M.L., and Henry, S.A.
S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase
compositions and methods for modulating lipid biosynthesis in
plants

JOURNAL Patent: WO 0168870-A 1 20-SEP-2001;

FEATURES University of Florida (US) ; Carnegie-Mellon University (US)

source

1. 2235
Location/Qualifiers

/organism="Spinacia oleracea"

/mol_type="unassigned DNA"

/db_xref="taxon:3562"

ORIGIN

Alignment Scores:

Pred. No.: 1,31e-195 Length: 2235

Score: 2204.00 Matches: 408

Percent Similarity: 93.64% Conservative: 34

Best Local Similarity: 86.44% Mismatches: 30

Query Match: 87.63% Indels: 0

DB: 6 Gaps: 0

US-10-031-331b-40 (1-473) x AX250551 (1-2235)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuApp 20
DB 317 CACTCTGTGTGATTTGACTGTGTGAGGCTATGATCTTGAATTCACAGCTTCAGATCTTCAC 376
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluGlyCysLeuLeu 40
DB 377 AAAGTGGAGCGACCTGAGGTACTTCCATGCTTCCACCTTATGAGGAAAGTGTGTCTTA 436
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyAlaGlyGlnVal 60
DB 437 GNACTCGGTGCTGTGTATGTGTCTTTTACTGTGGTAAATGGCCGAGAGAGTAGCCAGTGC 496
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLeuAsnGluValIleAsnGlyHisTyr 80
DB 497 ATCCCTCTGGATTTTCATTTGAGGTGTATTAAGAGAGATGAAGCATTAATGGGCATTC 556
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
DB 557 AAAAATGTGAAGTTTATGTGTCTGTATGTGACATCTCCAACTCTCAATTTCCCAAT 616
QY 101 SerLeuAspValIlePheSerAsnThrLeuLeuMetTyrLeuSerAspGluGluValGlu 120
DB 617 TCCGTGGATATCATTTCTCCAAATGGCTACTCATGTATCTTCTGTATGAAGAGGTTGAG 676
QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
DB 677 CGTCTGGTTCAGAGGATTTGAATGTTGAGCCAGGAGGATACATTTTCTTCAGAGAA 736
QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
DB 737 TCTTGTGTTTTCATCAATCAGAGATCATCAAGCGCAAAAGCAATCCCAACCTACCGTGAA 796
QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
DB 797 CCTAGGTTCTACACCAAGATCTTCAAGAGATGCCATATGCAAGATGATCTTGGGATCTCC 856
QY 181 TyrGluLeuSerLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200
DB 857 TATGAGTCTTCCCTTAATGGCTGCAATGATTTGAGCTTATGTCAAAAGCAGAGAGAT 916
QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
DB 917 CAGAACAGATAGCTGTGTTATGGCAGAGAAAGTTGATTCAGAGGATGACAGGGGTTCCAG 976

QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
DB 977 CGATTTCTTGGATTTCTAGTCAATCAAGATTTAAACAGCATACTGCGTTATGAGCGGTATTT 1036
QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
DB 1037 GGTCTGTGTATGTAGTACCGGAGGACTCCAAACCAACCAAGAGTTTGTATCAAGCTT 1096
QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
DB 1097 GACTTGAAGCTTGGCCAGAGGTCTTGAATGTGGTTTGGCATAGTGGAGGTGATTTT 1156
QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
DB 1157 TACATGGCAGAGAACTATGATGTTGAGGTTGTTGGAATGATCTCTCCATTAATATGAT 1216
QY 301 SerPheAlaLeuGluArgSerIleGlyLysLysCysAlaValGluPheGluValAlaAsp 320
DB 1217 TCTTTTGGCTTGGAGCGCTCAATTTGGCTTCAATGCTGTGTGAGTTTGGAGTGGCAGAT 1276
QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340
DB 1277 TGACCAAGAAAGATTAACCTTGAAACTCTTTTGAATGTCATCTACAGCGGTGATACCAT 1336
QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
DB 1337 CTGCATATTCAGGACAAACCTGCTTTATTATGATCTTCCACAAATGGTTGAACTGGA 1396
QY 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
DB 1397 GCAAGAGTCTTATTAGTACTTGTAGAGTGTGTGACCTTGTGACCTTCACTGCAATTTGCT 1456
QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGlyTyrGlyGlnMetLeu 400
DB 1457 GCATACATCAGGACAGGAGGATATGATCTCCAGATGTGAAGGCATATGGCAAGATGCT 1516
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
DB 1517 AAAGATCTCGATTCGTTGAGGTTATTTGCTGAGATAGAGCTGACCATTTCAATTCAGTT 1576
QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
DB 1577 CTGCAAGAGGAGTGTAGTCTCTTGAACAGGAGAGAGATGATCTTCAATTCATGATTTCTCT 1636
QY 441 GluGluAspTyrAsnAspIleValGlyTyrAsnAspLysLeuArgArgThrAlaLys 460
DB 1637 GAGGAGGATTTATAACGACATAGTTGATGTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1696
QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
DB 1697 GGTGAGCAACATGGGGTTTGTTCATTTGCCAAGAAA 1732

RESULT 5

AP237633

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

AP237633 2252 bp mRNA linear PLN 10-MAY-2000
Spinacia oleracea phosphoethanolamine N-methyltransferase (PEMT)

mRNA, complete cds.

AP237633

AP237633.1 GI:7407188

Spinacia oleracea (spinach)

Spinacia oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Amaranthaceae; Spinacia.

1 (bases 1 to 2252)

Nuccio, M.L., Ziemak, M.J., Henry, S.A., Weretilnyk, E.A. and

Hanson, A.D.

cDNA cloning of phosphoethanolamine N-methyltransferase from

spinach by complementation in *Schizosaccharomyces pombe* and

characterization of the recombinant enzyme

J. Biol. Chem. 275 (19), 14095-14101 (2000)

MEDLINE 20261526
FURNED 10799484
2 (bases 1 to 2252)
Ruccio, M.L. and Hanson, A.D.
Direct Submission
Submitted (21-FEB-2000) Horticultural Sciences, University of
Florida, P.O. Box 110690, Gainesville, FL 32611, USA
JOURNAL Location/Qualifiers
FEATURES
1..2252
/organism="Spinacia oleracea"
/mol_type="mRNA"
/db_xref="taxon:3562"
/tissue_type="sealined leaf"
1..2252
/gene="PEAMT"
234..1738
/gene="PEAMT"
/EC_number="2.1.1.103"
/function="catalyzes S-adenosylmethionine dependent
N-methylation of phosphoethanolamine,
phosphomonomethyl ethanolamine and
phosphodimethyl ethanolamine"
/codon_start=1
/product="phosphoethanolamine N-methyltransferase"
/protein_id="AA061950.1"
/db_xref="GI:7407189"
/translation="MAASMGVLOREVFVKYWIHSVDLTVEAMMLDSQSLDKVY
RPEVLSKLPYEGSKVLELGAGIGRTGELAKASQVIALOPIESVKWKNESLNGHYK
NVKPMCAVDTPSINSPNSVDIIIFSNLLATLSDEEYERLVERKMLKPGGVIPR
ESCFQSGDHRRKSNPTHPREFPIYKIFKCHMQDQSGNLSLGLCKICIGAYKS
KQNOQIQLMWQVSDSDKGFQFLDSQYKFNISLARYRVFGGVYSTGLLETKK
FVSKLDLPQKQVLDVCGIGGGDFMAENYDVRVVGIDLSINMISPALEKSLGKA
VEFEVADCTKKYDIPENSFOVYSRDTILHIDQKALPFRSFKMLKPGKVLISDYCK
AGTSAEFAAYIRQGVLDHVKAYGKMLKADGFVEVIAENRTDQFIQVLEKELDALE
QEKKDPIDDFSESDYNDIVDMKAKLVRTTEGQQWGLFIARKM"
ORIGIN
1.33e-195 Length: 2252
2204.00 Matches: 408
Score: 93.64% Conservative: 34
Percent Similarity: 86.44% Mismatches: 30
Best Local Similarity: 87.63% Indels: 0
Query Match: 8 Gaps: 0
DB: 0
US-10-031-331b-40 (1-473) x AF237633 (1-2252)
Qy 1 HistHValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
Db 317 CACTCTGTGTATTTGACCTGTTGAGGCTGATGATGCTTGTATTCACAGCTTCAGATTTTAC 376
Qy 21 LysGluAlaArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40
Db 377 AAATGGAGGACCTGAGTACTTCCATGCTTCCAGCTTATGAGGAGGAGTCTGTCTTA 436
Qy 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60
Db 437 GAACCTCGGTCGTGTATTTGCTGCTTACTGTTGATTTGGCCGAGGAGGAGTCCAGGTC 496
Qy 61 IleAlaLeuAspPheIleGlySerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
Db 497 ATCGCTCTGGATTTCAATGAGAGTGTTATAAAGAGAGATGAAAGCATAAATGGGCATTAC 556
Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProPheHis 100
Db 557 AAAATGTGAAGTTATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 616
Qy 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120
Db 617 TCGGTGATATATATTCCTCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyLysPheArgGlu 140

Db 677 CGTCTGGTGAAGAGATGTTGAATGGTTGAAGCCAGGAGATACATTTTCTTTCAGAGAA 736
Qy 141 SerCysPheHieGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
Db 737 TCTTGTGTTTTCATCAATCAGGAGATCACAAGGCGCAAGCAATCAACCCACTACCGTGA 796
Qy 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
Db 797 CTTAGGTTCTACACCAAGATCTTCAAGAGATGCCATATGCAAGATGATTTCTGGGAATCC 856
Qy 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200
Db 857 TAGAGCTCTCCCTAATTCGCTGCAATGATTTGGAGCTTATGTCAAAAGCAGAGAAT 916
Qy 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
Db 917 CAGAACCAAGATGCTGTTATGGCAGAAAGTTGATTCAGAGGATGACAGGGGTTCCAG 976
Qy 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
Db 977 CGATTCCTGGATTTCTAGTCAATACAGTTTAAACAGATCTGCGTTATGAGCGGTATTT 1036
Qy 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
Db 1037 GGTCTCTGTTATGTTAGTACCGGAGGACTCGAAGACACAGAGGTTTGTATCAAAGCTT 1096
Qy 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
Db 1097 GACTTGAAGCTGCGCCAGAGAGTCTTAGATGTGGGTGTGGCATAGGTGGAGGGTATTTT 1156
Qy 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
Db 1157 TACATGGCAGGAAGTATGATGTGAGTTGTGGATTTGATCTCTCCATTAAATGATTT 1216
Qy 301 SerPheAlaLeuGluArgSerIleGlyLysLysCysAlaValGluPheGluValAlaAsp 320
Db 1217 TCTTTGCTCTGAGCGCTCAATTTGGCTCAATTTGCTGTTGAGTTTGGAGTGGCAT 1276
Qy 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340
Db 1277 TGCACCAAGAGATTTACCTGTAAGACTCTTTTGTATGTCATCTACAGCGGTATACCAT 1336
Qy 341 LeuHisIleGlnAspLysProAlaLeuPheAspSerPheTyrLysTrpLeuLysProGly 360
Db 1337 CTGCATATTCAGGACAAACCTGCTTTATTTAGATCTCTTCACAAATGGTTGAACCTGGA 1396
Qy 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
Db 1397 GGCAAGATTTCTTATTTAGTACTACTGTAGAGTGTGGTACACCTTCAGCTGAATTTGCT 1456
Qy 381 AlaTyrIleLysGlnAspGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
Db 1457 GCATACATCGCCAGAGGGGATATGATCTCCAGTGTGAAGGCATATGCAAGATGCTT 1516
Qy 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
Db 1517 AAAGATGCTGGATTCGTTGAGGTTATTTGCTGAGATAGGACTGACAGTTTCATCAGTT 1576
Qy 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
Db 1577 CTGCAAGAGAACTAGATGCTCTTGAACAGGAGAGGATGACTTCATTCATGATTTCTCT 1636
Qy 441 GluGluAspTyrAsnAspIleValGlyLysTrpAsnAspLysLeuArgThrAlaLys 460
Db 1637 GAGGAGGATTTAAGCAGATAGTTGATGTTGGAGGCCAGGTTGTTGGAGGACTACAGAG 1696
Qy 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
Db 1697 GGTGAGCAAAATGGGGTTGTTGTTTCAATGCCAAGAA 1732
RESULT 6
AY133811
LOCUS AY133811 1459 bp mRNA linear PLN 18-SEP-2002

DEFINITION
Arabidopsis thaliana clone U1199 putative phosphoethanolamine N-methyltransferase (At1g48600) mRNA, complete cds.

ACCESSION
AY133811

VERSION
AY133811.1

KEYWORDS
FLI CDNA.

SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM
Arabidopsis thaliana

REFERENCE
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Arabidopsis Open Reading Frame (ORF) Clones

AUTHORS
Unpublished

TITLE
2 (bases 1 to 1459)

JOURNAL
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

REFERENCE
Submitted

AUTHORS
The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE
Submitted

JOURNAL
The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

COMMENT
The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Heuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

FEATURES
source
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.
Location/Qualifiers
1..1459
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U1199"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1..1459
/gene="At1g48600"
1..1428
/gene="At1g48600"
/codon_start=1
/evidence="experimental"
/product="putative phosphoethanolamine N-methyltransferase"
/protein_id="AAW91745.1"
/db_xref="GI:22136802"
/translations="MEHSSDLTVEAMMLDSKASLDKBERPEVLSPYPYEGKSVLEL

CAGTGRGTGELAQAGAEVIALDFEBSAIOKNSVNGHYKNIKPMCADVTSPDLKIDG
SIDLIFSNMLLMSDKVSELMABRMIGWPKGGYIFPFSKQSGSKSKSNPTHY
REPRFYTKVQBCQTRDASGNFELSMVCKICIGAYVKNQKQNOICMIWOKVSVEND
KDFPFLDNVOYKSSGILRYERFVGEGYVSTGGFTTKEFVAKMDLKPQOKVLDVGGC
IGGSDPYMAENFDVHVIGIDLSVNMISFALERAIGLKSVPVADACTTKTVPNSED
VIYSRDTILHIQOKPALFRFFFWLKPQCKVLITDYCRSAETPSPEFAEYLKORGYDL
HDQAYGQMLKADGPDQVIAEDRTDQVQVLRLELKEVEKEEFISDFSEEDYNDIV
GWSAKLERTASGEGKWLFIADKK"
misc_difference 282
/gene="At1g48600"
/note="compared to genomic sequence"
/replace="c"
1429..1459
/gene="At1g48600"
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 5.22e-184 Length: 1459
Score: 2077.00 Matches: 379
Percent Similarity: 90.06% Conservative: 47
Best Local Similarity: 80.13% Mismatches: 47
Query Match: 82.58% Indels: 0
DB: 8 Gaps: 0
US-10-031-331B-40 (1-473) x AY133811 (1-1459)
QY 1 HistThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
DB 7 CAITCTAGTGTGATTTGACTGTTGAAGCTATGATGCTTGAAGCTCTTAAGCTCTTGAATTTGAC 66
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluGlyLysCysLeuLeu 40
DB 67 AAGAGAGAGAGCTCTGAGGTACTCTCTTTATCCCAACCATATGAGAGGGAATCTGTGCTT 126
QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluLeuAlaGlyGlnVal 60
DB 127 GAACCTGGAGCTGTGATTTGCTGCTTTTCACTGGTGAATTTGGCTCAAAAGCTGGTGAAGTT 186
QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
DB 187 ATGCTCTTGAATTCATGAG 246
QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
DB 247 AAGAACATCAAGTTATGTGCTGTATGATGATATCCAGAGCTTGAAGAGTCAAGAGATGGA 306
QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120
DB 307 TCTATGACCTTGAATTTCTCAAACTGTTGCTCATGATATCTCTCTGATTAAGAGAGTGA 366
QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
DB 367 CTATGCGACAG 426
QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
DB 427 TCTTGTCTTCATCATCATCTGGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
DB 487 CCCAGATTTCTACACAAAGGTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 181 TyrGluLeuSerLeuSerCysLysCysLysGlyAlaTyrValArgAsnLysLysAsn 200
DB 547 TTGAGCTCTCTATGCTTGGCTGCAAAATGCTATGGGCTTATGTGAAGAGAGAGAGAGAG 606
QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspLeuGlyPheGln 220
DB 607 CAGAACTCAGATTTGCTGGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
DB 667 CGTTTCTTGACCAATGTTCAATACAGTCTAGTGGAGATCTTGGCTATGAGCGCTGCTTT 726

Db 85 |||||AAAGAGACGCTCTGAGTACTCTCTTTAAATCCACCATATGAGCGGAATCTGCTT 144
 Qy 41 GluLeuGlyValAlaGlyIleGlyValGlyPheThrGlyGluLeuAlaGlyValGlyVal 60
 Db 145 GAACTTGGAGCTGGTATTGGTCTGTTCTCACTGGTGAATGGCTCAAAAGGCTGGTGAAT 204
 Qy 61 IleAlaLeuAspPheIleGluSerAlaIleGlyValGluValIleAlaGlyHisTyr 80
 Db 205 ATCGCTCTGCTACTCTATCGAAGCGCAATTCAGAGAAGTGAAGTGTAAATGGGATTCAC 264
 Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
 Db 265 AAGAACATCAAGATTATGTGTGCTGATGTAAACATCTCCAGACTTGAATAAAGATGGA 324
 Qy 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluValGlu 120
 Db 325 TCTATCGACTTGATTTCTCAACTGGTGTCTCAATGATCTCTCTGATTAAGAGGTGGNA 384
 Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
 Db 385 CTAAATGGCAGAGAGAAATGATGGATGGTCAAGCCAGGGGATACATTTCTTCAGAGAA 444
 Qy 141 SerCysPheHisGlnSerGlyAspHisLysAlaGlySerAsnProThrHisTyrArgGlu 160
 Db 445 TCTTGCTTCCATCAATCTGGGACAGCAAGCGAAAGTCAAAACCCCACTCACTACCGTGA 504
 Qy 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 Db 505 CCCAGATTCTACACAAAGTTTCCAGGAATGTCAGACACGATGATGCTTCGGCAATTCAC 564
 Qy 181 TyrGluLeuSerLeuSerCysLysCysAlaGlyValAlaTyrValArgAsnLysLysAsn 200
 Db 565 TTTCAGCTCTCTATGTGGTGGTGCMAATGCAATGGGGCTTATGTGAAGAACAGAGAAAT 624
 Qy 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
 Db 625 CAGAATCAGATTCTCGATATGCAAAAGTCCAGCTGGAGATGACAGAGATTTCCAG 684
 Qy 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
 Db 685 GTTTCCTGGCAATGTTCAATACAGTCTAGTGGGATCTTGGCTATGACGGTGTCTTT 744
 Qy 241 GlyProGlyTyrValSerThrGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 Db 745 GCGGAAGATATGTGACACTGTGTGATTTGAGACAACTTAAGATTTGTGGCGAAGATG 804
 Qy 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
 Db 805 GACCTTAACCGGACGAAGTCCTAGATGTTGGTGTGGTATCGGTGGAGGTGACTTC 864
 Qy 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
 Db 865 TACATGGCTGAGAAATTCGATGTTTCATGTTGTGGAACTGCTGTGCTCAACATGATC 924
 Qy 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 Db 925 TCTTTCGACTGGAGCGGCGCATGACTCAATGCTCAGTGGATTTGAAGTGGCTGAT 984
 Qy 321 CysThrLysIleLeuSerProAsnSerPheAspValIleTyrSerArgAspThrIle 340
 Db 985 TGCACCAACCAAAACATATCCGATAATCTCTTGTATGTCATTTACAGCGGTGACATATT 1044
 Qy 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
 Db 1045 CTGCACATCCAGACAGCCAGCTCTATTTCAGACATTTCTCAATGCTGTAAACACAGG 1104
 Qy 361 GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
 Db 1105 GGTAAATGCTCATCATGACTATTCAGAGCTGCTGAACCTGCTCTCTGATTCGCA 1164
 Qy 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400

Db 1165 GAGTACATAAAACAAAGAGGATATGATCTACATGNTGTTCAAGCTTACGACAGATGCTG 1224
 Qy 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
 Db 1225 AAAGACGAGCGCTTTGACGAGCTTATCGCTGAGGACCGTACTGATCAGTTGTACAGTC 1284
 Qy 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
 Db 1285 CTCAGCGTGAATTAGAAAAAGTGGAGAAAGAAAGAAAGATTCATCAGCGACTTCTCA 1344
 Qy 441 GluGluAspTyrAsnAspIleValGlyCysTrpAsnAspLysLeuArgArgThrAlaLys 460
 Db 1345 GAAGAGGATTCATGATGATGTTGGAGGATGTCGGAAAGCTTGAAGGAGTGCATCT 1404
 Qy 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLysLys 473
 Db 1405 GGTGAACAGAAATGGGATTTATTCATAGCCGACAGAAAG 1443

RESULT 8 AY063866

LOCUS 1660 bp mRNA linear PLN 18-SEP-2002
 DEFINITION Arabidopsis thaliana putative phosphoethanolamine
 N-methyltransferase (At1g48600) mRNA, complete cds.

ACCESSION AY063866
 VERSION AY063866.1 GI:17380783
 KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1660) Arabidopsis thaliana (thale cress); Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE AUTHORS

Yanada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C.J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Sakurai, T., Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 Unpublished

TITLE JOURNAL REFERENCE AUTHORS

2 (bases 1 to 1660)
 Yanada, K., Banh, J., Banno, P., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission
 Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE JOURNAL COMMENT

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Banno, P., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kawai, J., Koesema, E., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis


```

genome submitted to GenBank.
Location/Qualifiers
1..1660
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon.3702"
/chromosome="1"
/clone="RAF106-86-D12 (R11999)"
/ecotype="Columbia"
/notes="This clone is in a modified pBluescript vector
(Lambda ZAP) as a XhoI/SatI insert."
1..1660
/gene="At1g48600"
1..34
/gene="At1g48600"
35..1462
/gene="At1g48600"
/codon_start=1
/evidence=experimental
/product="putative phosphoethanolamine
N-methyltransferase"
/protein_id="AA136222.1"
/db_xref="GI:17380784"
/translators="MEHSDLTVEAMMLDSKSLDKBERPEVLSLIPPEKGSVLEL
GAGRTPTGLAOKAGEVIALDFISAIQKNSVNGHYKFKFMCADVTSPLDKIKDG
SIDLFNNLLAYLSKDEGLBMAIRMGVPGGYFFRESCFHOSGDSKRSKNPHY
REPRFTYVQECOTRDASCNSPELSMGCKICGAVNKNQNOQIWMOKVSDGVS
KDFORFLDNVQKSSGILRYRVEGSGYVGTGPTTEPKVAKNDLKPQKVLVDGSD
IGGDFTMENFDVHVVGDLSVNMISPALEAIGLKSVEFEVADCTTKYTPNSGD
VIYSRDTLHIDQKPALEFRTPFKWLPAGKVLITDTCRAETPSPEFARYIKQGYDL
HDVQYQGMKDGDFDVIABDRDQFVQVLRRELEKVEKEEFISFSEEDINDIV
GGWAKLERTASGSKWGLFIADKK"
misc_difference 316
/gene="At1g48600"
/note="compared to genomic sequence"
/replace="t"
1463..1660
/gene="At1g48600"

3' UTR
1463..1660
/gene="At1g48600"

ORIGIN
Alignment Scores:
Pred. No.: 6,198-194 Length: 1660
Score: 2077.00 Matches: 379
Percent Similarity: 90.06% Conservative: 47
Best Local Similarity: 80.13% Mismatches: 47
Query Match: 82.58% Indels: 0
DB: Gaps: 0

US-10-031-331B-40 (1-473) x AY063866 (1-1660)
Qy 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuApp 20
Db 41 CATCTAGTGAATTTGACTGTTGAAGCTATGATGCTTGACTCTAAAGCTTCTGATCTTGAC 100
Qy 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGlyLysCysLeuLeu 40
Db 101 AAGAGAGACGCTCTGAGGTACTCTCTTAATCCACCATATGAAGGAAATCTGTGCTT 160
Qy 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGlyLysAlaGlyGlnVal 60
Db 161 GAATCTGGAGCTGGTATGTTGCTTCACTGGTGAATGGCTCAAAAGGCTGGTGAGTT 220
Qy 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
Db 221 ATCGCTCTTGACTTCATCCAAAGCGGATTCGAGAAATGAAAGTGTAAATGGGCATTAC 280
Qy 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
Db 281 AAGACATCAAGTTATGTGCTGATGTATACATCCCGACCTTGAAATCAAGATGGA 340
Qy 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluGluValGlu 120
Db 341 TCTATCGACTTGAATTTCTCAACTGGTGTCTCATGTATCTCTCTGATAAAGAGGTGGA 400

```

```

Qy 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
Db 401 CTAATGCGACAGAGAAATGATGGATGGGTCAAGCCAGCGGGGATACATATTTCTTCAGAGAA 460
Qy 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
Db 461 TCTTGCTTCCATCAATCTGGGACAGCAAGCGAAAGTCAAAACCCCACTCACTACCGTGA 520
Qy 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
Db 521 CCCAGATTCTACAAAGGTTTTCAGGAATGTACAGACAGCTGATGCTCTTCGGCAATTC 580
Qy 181 TyrGluLeuSerLeuLeuSerCysLysCysIleGlyAlaTyrValArgAsnLysLysAsn 200
Db 581 TTTGAGCTCTATGTTGGCTGCAAAATGCATTTGGGGCTTATGTGAGAGACACAGAGAAAT 640
Qy 201 GlnAsnGlnIleSerTrpLeuTyrGlnLysValAspSerLysAspAspLysGlyPheGln 220
Db 641 CAGATTCAGATTCTGCTGATATGCGAANAAGTCAGCGTGGAGATGACACAGGATTTCCAG 700
Qy 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
Db 701 CGTTTCTTGGACAATGTTCAATCAAGTCTAGTGGGATCTTCGCTATGAGCGGTCTCTT 760
Qy 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
Db 761 GGGGAGGATATGTGACACTCGTGCATTTGACACAACTAAAGAAATTTGTGGCGAAGATG 820
Qy 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyValAspPhe 280
Db 821 GACCTTAACCGGACAGAAAGTCTAGATGTTGGTGTGATCGTGGAGGTGATCTTC 880
Qy 281 TyrMetAlaGluThrPheAspValGluValGlyPheAspLeuSerValAsnMetIle 300
Db 881 TACATGGCTGAGAAATTTGATGTTGTTGGTATCGATCTGTGCGGTCAACATGATC 940
Qy 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
Db 941 TCTTTCGCACTGGAGCGGGCCATTTGCAATGCTCAGTCGAGTTTGAAGTTCGCTGAT 1000
Qy 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340
Db 1001 TGCACCAACCAACATATCCGATATTCCTTTGATGTATTTACAGCGCTGACACTATT 1060
Qy 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
Db 1061 CTGCACATCCCAAGCAAGCCAGCTCTATTGAGGACATTTCTCAAGTGGGCTTAAACCGAG 1120
Qy 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
Db 1121 GGTAAAGTTCTCATCTGACTATTTGCAAGAGTGTGCAAACTCGTCTCTCGAATTCGCA 1180
Qy 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
Db 1181 GAGTACATAAAACAAAGAGGATATGATCTCATGATGTTTCAAGCTTACGAGACAGATGCTG 1240
Qy 401 LysAspAlaGlyPheValAspValIleAlaGluAspArgThrGluGlnPheIleArgVal 420
Db 1241 AAAGCGCAGGCTTTGACGACGTTATCGTGGAGCGCTACTGATCAGTTTGTACAGTTC 1300
Qy 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
Db 1301 CTGAGCGGTGAATTAAGAAAGTGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1360
Qy 441 GluGluAspTyrAsnAspIleValGlyTyrPheAsnAspLysLeuArgArgThrAlaLys 460
Db 1361 GAAGAGGATTAACATGACATTTGTTGGAGGATGTTGGCGCAAGCTTGAAGAGGATGATCT 1420
Qy 461 GlyGluArgTyrGlyLeuPheValAlaLysLysLys 473
Db 1421 GGTGACAGAGAAATGGGATTTATCATGCCGACAGAG 1459

```

RESULT 9

AF328858
 LOCUS
 DEFINITION Lycopersicon esculentum phosphoethanolamine N-methyltransferase
 mRNA, complete cds.
 ACCESSION AF328858
 VERSION AF328858.1 GI:12584942
 KEYWORDS Lycopersicon esculentum (tomato)
 SOURCE Lycopersicon esculentum
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 1794)
 AUTHORS Parani, M. and Parida, A.
 TITLE Characterization of a cDNA for phosphoethanolamine N-methyltransferase
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1794)
 AUTHORS Parani, M. and Parida, A.
 TITLE Direct Submission
 JOURNAL Submitted (13-DEC-2000) Plant Molecular Biology, M. S. Swaminathan Research Foundation, III Cross Street, Chennai, Tamil Nadu 600 113, India

FEATURES

source

1..1794
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 106..11581
 /codon_start=1
 /product="phosphoethanolamine N-methyltransferase"
 /protein_id="AA059894.1"
 /db_xref="GI:12584943"
 /translations="MAAAGQREEVQKSVMIETAEITVEMWLDKADLKERPE
 VLSLPPEKTVLEAGIGFTGELAGKAGQLALDPIEGAIKQENINIGHTKQVK
 PMCADVTPNLIIPESVQLIFSNLLATLSDEEVKALVERWVILKVGHHIPRES
 PHOGDHKRNKTHTRDPSFTFKVFRCHYNAGDGKFSLSAGCKICIGAYRNKN
 QNLCIKWQTSKDMERFQDLTVQKSGILRYERFQGVFVSGLETKFEVA
 MLDLRQKVDVGGGIGNFMAEKYDVRVGVGIDLSINMISFALRAIGLCAVEF
 EVADCTKPTDCTFVTSRDTLIHQDKPALFRPYKWLKRPQGLVLSIDYCKRAGP
 ASKEFEGYIKQRYDLDHVEAYQMLDAGPHEVVAEDTEQFIKVLQKSLDTVEKER
 ESPTHEFSEQDYNEIVGGWKSILIRSSGQERWGLFIKKK"

ORIGIN

Alignment Scores:
 Pred. No.: 1,38e-182 Length: 1794
 Score: 2063.00 Matches: 379
 Percent Similarity: 89.85% Conservative: 46
 Best Local Similarity: 80.13% Mismatches: 48
 Query Match: 82.03% Indels: 0
 DB: 8 Gaps: 0

US-10-031-331B-40 (1-473) x AF328858 (1-1794)

QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 DB 160 CACACTGCGGAACTCACTGTGAGGCGCATGCTCGACTCTAAAGCAGCGGATCTCGAC 219
 QY 21 LysGluGlnArgProGluIleLeuSerMetLeuProLeuGluGlyIleCysLeuLeu 40
 DB 220 AAAGAAGAGAGACCTGAGGTGCTCTCTGCTTCGCGCCGATGAGGAGAAACTGTGTG 279
 QY 41 GluLeuGlyAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluAlaGlyGlnVal 60
 DB 280 GAACCTGGCTGAGGATATGCGCATCTCTGCTGATGATAGCCAGAAAGCTGGACGCTT 339
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysAsnGluValIleAsnGlyHisTyr 80
 DB 340 ATAGCTGGACTTATTTGAAGCGCAATTAAGCAAGATGAATATTAATGCGCACTAT 399
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100

DB 400 AAGATGTCAAGTTCATGTGCTGATGACTTCCCGAATTTGATTTTTTTCACCGAA 459
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuLeuMetTyrLeuSerAspGluValGlu 120
 DB 460 TGGTGGACTTGAATATTTCTTAAGTCTACTGATGATATCTTCTGATGAAGAGTTAAG 519
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyIlePhePheArgGlu 140
 DB 520 GCTCTTGTTCAGAGAAATGGTTATATATGTTGAAAGTCGTGGTGCACATATTTTCAGAG 579
 QY 141 SerCysPheHisGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu 160
 DB 580 TCACTCTTCATCATCATCAGGAGACCAAGCGAAGAAACCAACCCACCATATTCAGAG 639
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLysLeuGlnAspGlySerGlyAsnSer 180
 DB 640 CCTAGCTTTTATACAAAGGTTTTTTCAGAGAAATGTCAGTAAATGCTGGTGAATCA 699
 QY 181 TyrGluLeuSerLeuLeuSerCysLysCysValIleGlyAlaTyrValArgAsnLysAsn 200
 DB 700 TTTGAACCTTCTCTTGGTGGTTCGAAGTGCATTTGGAGCTTATGTGAGACAAAGAAAT 759
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAspAspLysGlyPheGln 220
 DB 760 CAAATCAGATTTGCTGGACGTGGCAGAGGTTACTTCTTAAGGATGACATGGAATTCAG 819
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuAlaTyrGluArgValPhe 240
 DB 820 CGTTTCTTGGATCTGTTTCAGTACAAAGTGTAGTGGCATCTAGCTATGACAGGTCTTT 879
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 DB 880 GGACAGGTTTGTGTAGCAGGAGGACTTGAACATACAAAGAAATTTGTAGCCATGTTG 939
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
 DB 940 GATCTTCGCGTGGCCAAAAGTCTTGTATGTTGGCTGGCATTTGGAGGAGGTAACCTTC 999
 QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
 DB 1000 TACATGCTGAAAGATGATGTTTCATGTTGTGGCATTTGCTGTCTATCAATGATA 1059
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 DB 1060 TCCTTTTCTCTGAGCGTGTCTATTTGCTCAAGTGTGCTGTTGAAATTTGAGGTGCTGAT 1119
 QY 321 CysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThrIle 340
 DB 1120 TGCACCAAGAAACATATCTGATTTGATGATGATGATGATGATGATGATGATGATGAT 1179
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
 DB 1180 CTTCACTCAGGACCAACCTGCATCTTCATGATCTTTCTACAAAGTGGCTGAGACCAAG 1239
 QY 361 GlyLysValLeuLysSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
 DB 1240 GGCAAGGTTCTTAAAGTGATTTACTGCAAAAGAGCTGCAACAGCATCAAAAGAAATTCGA 1299
 QY 381 AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
 DB 1300 GGTATATTAAAGCAAGGGGTATGATTTACATGATGTTGAAGCATATGGCAGATGCTC 1359
 QY 401 LysAspNlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
 DB 1360 AGAGATCGCGTTTTTCATGATTTGTGGCCGAGGATGCTTACTGAACAGTTCAATTAAGGT 1419
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
 DB 1420 CTCGAAAGGAGGTAGATATCTGTGGAGGAGGAGGAGGATCTTTTATTCATGAGTCTCT 1479
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgThrAlaLys 460
 DB 1480 GAACAAGACTATTAATGAATAGTTGGAGGTTGGAGGTTGGAGGTTCAATCAGGAGTTCACT 1539

QY	461	Glycylglutarylarginylglutamate	1870 bp	mRNA	linear	PLN 04-MAR-2003
DB	1540	GCTGACGACGAGGCTGGGCTGTTTCATTCGACGAGGAA	1578			
RESULT 10						
AB090883						
LOCUS						
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REFERENCE						
TITLE						
JOURNAL						
FEATURES						
source						
CDS						
ORIGIN						
Alignment Scores:						
Pred. No.:	7,778-181	Length:	1870			
Score:	2044.50	Matches:	374			
Percent Similarity:	89.66%	Conservative:	51			
Best Local Similarity:	78.90%	Mismatches:	48			
Query Match:	81.29%	Indels:	1			
DB:	8	Gaps:	1			
US-10-031-331B-40 (1-473) x AB090883 (1-1870)						
QY	1	HistHrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp	20			
DB	248	CACCTGTTGATCTTACTGTTGAATCTATGATCTGATCTGATCTGATCTGATCTGAT	307			
QY	21	LysGluGluArgProGluIleLeuSerMetLeuProLeuGluGluGluGluGluGluGlu	40			
DB	308	AAACAGAAAGGCTGAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	367			
QY	41	GluLeuGluAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	60			
DB	1448	CTTAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1507			
DB	368	GAACTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	427			
QY	61	IleAlaLeuAspPheIleGluSerAlaIleIleIleIleIleIleIleIleIleIleIleIleIle	80			
DB	428	ATAGCCCTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT	487			
QY	81	LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis	100			
DB	488	AAAAATGTCAGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	547			
QY	101	SerLeuAspValIlePheSerAsnTrpLeuMetCysLeuSerAspGluGluValGlu	120			
DB	548	TCCATGCAATTTGATTTCTCAATTTGATTTCTCAATTTGATTTCTCAATTTGATTTCTCAAT	607			
QY	121	AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu	140			
DB	608	GATATTCGGAAGGTTTTTGAATTCGGTGAAGGTGGAGGACATATTTTCTTCAGAGAG	667			
QY	141	SerCysPheHISGlnSerGlyAspHisLysArgLysSerAsnProThrHisTyrArgGlu	160			
DB	668	TCATGCTTTCCACCATCTGGACCATAGCGAACAACCCCTACGACCTACCGGCA	727			
QY	161	ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer	180			
DB	728	CCTAGATTTCTACCAAGGCTTTCAAGCAATGTCATGACGATGCAATCTGGGAAATTC	787			
QY	181	TyrGluLeuSerLeuLeuSerCysLysCysValIleGlyAlaTyrValArgAsnLysLysAsn	200			
DB	788	TATGAATCTCTGCTAAATTCGATCAAAATGCAATTTGGAGCTTATGTCGCGACGAAGAAAT	847			
QY	201	GlnAsnGlnIleSerTrpLeuTrpGlnLysVal---AspSerLysAspAspLysGlyPhe	219			
DB	848	CAACCCAGATCTCTGATATGCGAGAGTGGCGAGTTCAGAGGACGATAGAGGCTTC	907			
QY	220	GlnArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgVal	239			
DB	908	CAGCAATCTTGGATAATGCGCAGTACAAAGCAGTGGCATTTCTACGCTATGAACGATA	967			
QY	240	PheGlyProGlyTyrValSerThrGlyGlyTyrGlnThrThrLysGluPheValSerMet	259			
DB	968	TTTGACCCAGGTTTGTGACGACAGGCGGAGTAGATCCCAAAAGAAATTTGGTGCATTC	1027			
QY	260	LeuAspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAsp	279			
DB	1028	CTTGATTTAAAGCCCGCCGACAGAAATTTGGATGTTGGATGTTGGATGTTGGATGTTGGAT	1087			
QY	280	PheTyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMet	299			
DB	1088	TTCTATATGCGAGATAACTTCGATGTTGATTTGATTTGATTTGATTTGATTTGATTTGAT	1147			
QY	300	IleSerPheAlaLeuArgSerIleGlyLysLysCysAlaValGluPheGluValAla	319			
DB	1148	ATAGCTTTTGGCTCTTGAAGATCAATTTGGCTCCATTTGTTCAATTTGATTTGATTTGAT	1207			
QY	320	AspCysThrLysIleAsnTyrProAspAsnSerPheAspValIleTyrSerArgAspThr	339			
DB	1208	GACTGTACTTAAAGAAATCATATTCGACCAATTTGTTGATTTGATTTGATTTGATTTGAT	1267			
QY	340	IleLeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysPro	359			
DB	1268	ATCTCTCATATCCAGGATAAATCTGCTTTGTTCCGAACCTTCTACAGTGGTTGAAACCA	1327			
QY	360	GlyGlyValValLeuIleSerAspTyrCysLysValAlaGlyProProSerProGluPhe	379			
DB	1328	GGTGGCAAGATTTCTCATGATGATTTGCTGCGCAAAATCTGCAAAACCTCTGAGATTTT	1387			
QY	380	AlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMet	399			
DB	1388	GCAGAGTACATCAAGCAAGAGGATATGATCTCCATGATGTTGAAATTTATGGCAAGATG	1447			
QY	400	LeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArg	419			
DB	1448	CTTAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1507			

261 AspLeuLysProGluGlnValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
 946 GATCTGAACCTGGCCAAAGTCTTAGAGTGTGGTCCGATAGAGGAGGGGACTTC 1005
 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValanMetIle 300
 1006 TACATGGCTGAGAACTTTGACGTGGATGTTGGGCATTTATCTATCTGTAACATGATC 1065
 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 1066 TCITTTGGCTTCAACACGCAATAGGACTCAATCTCTGTAGAAATTCGAAGTAGCTGAT 1125
 321 CysThrLysIleAsnTyrProPheAsnSerPheAspValIleTyrSerArgAspThrIle 340
 1126 TGCACCAAGAGGAGTATCTGTATACACCTTTGATTTGATTTATATGACGAGACCAT 1185
 341 LeuHisIle----- 343
 1186 CTACATATCCAAAGTATAGCTTCGGCTATTTCTTCTTCCCTATGACTTTAATTTGGTAT 1245
 344 -----GlnAspLysProAlaLeuPheArgSerPheTyr 354
 1246 CCAATTTTGAATCCCTGCTTACTCTCAGGACCAAGCCAGCATTTCTTCAAGATTTCTAC 1305
 355 LysTyrLeuLysProGlyGlyValLeuLysValLeuLysSerAspTyrCysLysAlaGlyPro 374
 1306 AAATGGTGTGAACCGGAGGGAAGTTCATCATCTGATCTTCTGAGAGCCCAAAACC 1365
 375 ProSerProGluPheAlaAlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLys 394
 1366 CCATCTCCAGACTTTGCAATCTACATCAAGAACGAGGTATGATCTTCATGATGTACAA 1425
 395 GluTyrGlyGlnMetLeuLysAspAlaGlyPheValAspValLeuAlaGluAspArgThr 414
 1426 GCATACGCTCAGATGCTGAGAGTCTGTTTCGAGGAGTAAATCGGAGGATAGAAC 1485
 415 GluGlnPheIleArgValLeuArgLysGluLeuGluThrValGluLysGluLysAspVal 434
 1486 GATCAGTTCATGAAGTCTTCAAGCGGAACTGGTGCAGTGGAGAGGAGGAGGAGAA 1545
 435 PheIleSerAspPheSerGluGluAspTyrAsnAspIleValGlyLysTrpAsnAspLys 454
 1546 TTCATCAGTACTTCTCGAAGAGGATACGAGGATATATAGCGGGTGGAGTCAAG 1605
 455 LeuArgArgThrAlaLysGlyGluGluArgTyrGlyLeuPheValAlaLysLys 472
 1606 CTACTTAGGAGCTCAAGTGGTGGAGGAGTGGGGTTTGTTCATCCCAAGAGA 1659

RESULT 12
 AV091683 1476 bp mRNA linear PLN 13-APR-2002
 DEFINITION Arabidopsis thaliana AT3g18000/MB5_22 mRNA, complete cds.
 ACCESSION AV091683
 VERSION AV091683.1 GI:20147130
 KEYWORDS FLI CDNA
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1476)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 AUTHORS Bower, R., Chen, H., Kim, C.J., Meyers, M.C., Shinn, P., Ban, J.,
 Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
 Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,
 Theologis, A., and Ecker, J.R.

TITLE Arabidopsis ORE clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1476)
 AUTHORS Cheuk, R., Chen, H., Kim, C.J., Meyers, M.C., Shinn, P., Ban, J.,

Bower, R., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
 Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
 Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
 Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
 Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,
 Theologis, A., and Ecker, J.R.

Direct Submission
 Submitted (21-MAR-2002) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
 Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGE (SSP) Consortium members carried out the
 sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H.,
 Kim, C.J., Meyers, M.C., Shinn, P., Ban, J., Bower, R., Chan, M.M.,
 Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Jones, T.,
 Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C.,
 Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Davis, R.W.,
 Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES
 source
 1..1476
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="U12246"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51"
 1..1476
 /note="putative methyltransferase"
 /codon_start=1
 /product="AT3G18000/MB5_22"
 /protein_id="AA010282.1"
 /db_xref="GI:20147131"
 /translation="MAAYEEEDIDQNYWIEHSDLTVEAMMLDSRADSLDKBRPE
 VLSLPPVEGKSVLEAGTGRFTGAGELIALDFIDVINKNESINGHYKQVK
 PNCADVTSPDLKIDGSLDIFSNWLLVLSKVELLAERVMGVIKVGVIYPPFSC
 PHOSGDSKRSNPTHYRPPYKVFQECOTDAAGNSFELSMIGCTGATYKNNKN
 QNQCIMQKVSSENDERFORPLDQYKSSGILRVYVGGQFVSTGLGTTKEFVE
 KNNLAKPCQNLVDCGCGGDFYMAKPDVHVVDLVSNMISPALERAIGLSCSVF
 EVADCTTHYFDSFVYSRDTILHILQKPALEFPKPKAGKGVLSIDYCRSPKT
 PSASFSEYIKQGYLDHVDQATGOMLKDAGFTDVIADTDQPMQVLEBLDRVEKEK
 EKFI DFSKEDYDDIDVGWKNKLERCASDDEKQWGLFIANNK"

CDS
 1..1476
 /note="putative methyltransferase"
 /codon_start=1
 /product="AT3G18000/MB5_22"
 /protein_id="AA010282.1"
 /db_xref="GI:20147131"
 /translation="MAAYEEEDIDQNYWIEHSDLTVEAMMLDSRADSLDKBRPE
 VLSLPPVEGKSVLEAGTGRFTGAGELIALDFIDVINKNESINGHYKQVK
 PNCADVTSPDLKIDGSLDIFSNWLLVLSKVELLAERVMGVIKVGVIYPPFSC
 PHOSGDSKRSNPTHYRPPYKVFQECOTDAAGNSFELSMIGCTGATYKNNKN
 QNQCIMQKVSSENDERFORPLDQYKSSGILRVYVGGQFVSTGLGTTKEFVE
 KNNLAKPCQNLVDCGCGGDFYMAKPDVHVVDLVSNMISPALERAIGLSCSVF
 EVADCTTHYFDSFVYSRDTILHILQKPALEFPKPKAGKGVLSIDYCRSPKT
 PSASFSEYIKQGYLDHVDQATGOMLKDAGFTDVIADTDQPMQVLEBLDRVEKEK
 EKFI DFSKEDYDDIDVGWKNKLERCASDDEKQWGLFIANNK"

ORIGIN
 Alignment Scores:
 Pred. No.: 7,57e-178 Length: 1476
 Score: 2011.00 Matches: 367
 Percent Similarity: 88.56% Conservative: 51
 Best Local Similarity: 77.75% Mismatches: 54
 Query Match: 79.96% Indels: 0
 DB: Gaps: 0
 US-10-031-331b-40 (1-473) x AY091683 (1-1476)

Qy 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
 Db 55 CATTCGCTGATCTGACTGCTGAGCTATGATGCTTGCATCGAGGCTTCTGATCTCGAC 114
 Qy 21 TybGluGluArgProGluLysLeuSerMetLeuProLysGluGlyLysCysLeuLeu 40
 Db 115 AAGGAGAGAGCTCTGAGTACTCTCTTGTCTCTCCATATGAGGCAATCAGTGTG 174

QY 41 GluLeuGlyAlaGlyIleGlyAArgpThrGlyGluLeuAlaGluLysAlaGlyGlnVal 60
 Db 175 GAACCTGGAGCTGGTATTGGTCTGTTCACTGGTGAATAGCTCAAAAGGCTGGTGAATC 234
 QY 61 IleAlaLeuAspPheIleGluSerAlaIleLysLysAsnGluValIleAsnGlyHisTyr 80
 Db 235 ATTGCTCTGACTTCATTGTAACGTTTCAAGAGAGATGAAGATATCAATGGGCACTTAC 294
 QY 81 LysAsnValLysPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
 Db 295 AAGAAATGTCAGTTTATGTGCTGATGTTACATCCCTGACCTCAAGATCACTGATGGA 354
 QY 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120
 Db 355 TCTCTGACTTGATTCTCCAACTGGCTCTCATGTATCTTCTGCAAGAGAGGTGGAG 414
 QY 121 AsnLeuValGluArgMetLeuLysTrpLeuLysProGlyGlyTyrIlePhePheArgGlu 140
 Db 415 CTTTGTGGCAGAAAGGATGGTCGGTTGGATCAAGGTTGGAGGATACATTTCTTCCTGGAA 474
 QY 141 SerCysPheHisGlnSerGlyAspHisLysAlaGlyLysSerAsnProThrHisTyrArgGlu 160
 Db 475 TCTTGTCTCCAACTCAGGGACAGTAAAGCGGAATCCAACTCACTACCTGGAA 534
 QY 161 ProArgPheTyrThrLysAlaPheLysGluCysHisLeuGlnAspGlySerGlyAsnSer 180
 Db 535 CCCGCTTCTATCCAGGTCTTTCAGAGGTCTCAGACTCGGGATGCTGCTGGAAATTC 594
 QY 181 TyrGluLeuSerLeuSerCysLysCysLysGlyAlaTyrValArgAsnLysLysAsn 200
 Db 595 TTTGAGCTCTCTATGATCGGATCGAGTCAAGTCAATGGAGCTTATGTCAAGAACCAAGAAAT 654
 QY 201 GlnAsnGlnIleSerTrpLeuTrpGlnLysValAspSerLysAsnAspLysGlyPheGln 220
 Db 655 CAGAAATCAGATTGTTGGATATGGCAGAGGTCACTCAGAAATGACAGAGGCTTCCAA 714
 QY 221 ArgPheLeuAspThrSerGlnTyrLysCysAsnSerIleLeuArgTyrGluArgValPhe 240
 Db 715 GCTTCTTGCAACATGTCACATACAAATCCAGTGAATCCATGGAATCTGACGTATGAGCGTGTCTTT 774
 QY 241 GlyProGlyTyrValSerThrGlyGlyTyrGluThrThrLysGluPheValSerMetLeu 260
 Db 775 GGCCAGAGGTTTGTGAGCACTGTGCACTGTGAGACCAACCAAGAAATTTGTGGAGAAATG 834
 QY 261 AspLeuLysProGlyGlnLysValLeuAspValGlyCysGlyIleGlyGlyAspPhe 280
 Db 835 AATCTGAAACACGACAGCAAGATCTTAGATGTGGGTGGCATTTGGTGGAGGTGACTTC 894
 QY 281 TyrMetAlaGluThrPheAspValGluValValGlyPheAspLeuSerValAsnMetIle 300
 Db 895 TACATGCTGAGAAATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 954
 QY 301 SerPheAlaLeuGluArgSerIleGlyLeuLysCysAlaValGluPheGluValAlaAsp 320
 Db 955 TCTTTCGATGGAAGCTGCTATTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
 QY 321 CysThrLysIleAsnTyrProAsnSerPheAspValIleTyrSerArgAspThrIle 340
 Db 1015 TGACACCAAAACACATACCAGATAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1074
 QY 341 LeuHisIleGlnAspLysProAlaLeuPheArgSerPheTyrLysTrpLeuLysProGly 360
 Db 1075 CTGCACATCCAGACCAACAGCTGTTGTTAGGACTTCTTCAATGGCTAAACCGGGA 1134
 QY 361 GlyLysValLeuIleSerAspTyrCysLysLysAlaGlyProProSerProGluPheAla 380
 Db 1135 GGTAAAGTCTCATCAGGACTACTGTAGAGCCCAAAATCCATCTCTGCTGAGTTTCA 1194
 ...-381-AlaTyrIleLysGlnArgGlyTyrAspLeuHisAspValLysGluTyrGlyGlnMetLeu 400
 Db 1195 GAGTACATCAACAGAGAGATATGATCTCCATGACGTTCAAGCTTATGACAGATGCTA 1254

QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
 Db 1255 AAGACCTCGCTTCACTGATGTGATCGCAGAGGACCGTACTGATCAGTTTATGCAAGTC 1314
 QY 421 LeuArgLysGluLeuGluThrValGluLysGluLysAspValPheIleSerAspPheSer 440
 Db 1315 CTGAACCTGAATTAGACAGGCTGGAGAAAGAAAGAAATTCATCTCCGACTCTCTCC 1374
 QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspLysLeuArgArgThrAlaLys 460
 Db 1375 AAGAGGATACGATGACATCTGTTGGAGGTGAGAGTCAAGCTCGAGAGGTGTCATCG 1434
 QY 461 GlyGluGlnArgTrpGlyLeuPheValAlaLysLys 472
 Db 1435 GATGAGCAGAAATGGGGACTTTTTCATCGCCAAACAG 1470

AV058175 1770 bp mRNA linear PLN 04-NOV-2001
 Arabidopsis thaliana AT3g18000/MEB5_22 mRNA, complete cds.

AV058175
 DEFINITION Arabidopsis thaliana (thale cress)
 ACCESSION AV058175
 VERSION AV058175.1
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 (bases 1 to 1770)
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J., Bower, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Arabidopsis cdna clones
 Unpublished

2 (bases 1 to 1770)
 Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Ban, J., Bower, L., Carninci, P., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Direct Submission

TITLE Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory
 JOURNAL (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length CDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, RPEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Ban, J., Bower, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Kamiya, A., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Nguyen, M., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
 Location/Qualifiers
 1..1770 source


```

QY 361 GlyIysValLeuIleSerAspTyrCysValYsaAlaGlyProProSerProGluPheAla 380
Db 1275 GGTAAAGTTCATCATGACGACTACTGTAGAGCCCAAACTCCATCTGCTGAGTTTCA 1334
QY 381 AlaTyrIleYsaGlnAsgGlyTyrAspLeuHisAspValYsaGlyGlyGluMetLeu 400
Db 1335 GAGTACATCAACAGAGAGATATGATCTCCATGAGTTCAAGCTTATGACAGATGCTA 1394
QY 401 LysAspAlaGlyPheValAspValLeuAlaGluAspArgThrGluGlnPheIleArgVal 420
Db 1395 AAAGAGCTGGCTTCACTGATGTGATCGCAGAGACCGTACTGATCAGTTTATGCAATC 1454
QY 421 LeuArgYsaGluLeuThrValGluYsaGluYsaAspValPheIleSerAspPheSer 440
Db 1455 CTGAACCTGATATGACAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1514
QY 441 GluGluAspTyrAsnAspIleValGlyGlyTyrAsnAspYsaGluArgThrAlaYsa 460
Db 1515 AAAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574
QY 461 GlyGluGlnArgTyrGlyLeuPheValAlaYsaYsa 472
Db 1575 GATGAGCAGAAATGGGACTTTTCATGCGCAACAG 1610

RESULT 15
AF197940 1878 bp mRNA linear PLN 19-DEC-2000
LOCUS Arabidopsis thaliana SAM:phospho-ethanolamine N-methyltransferase
DEFINITION (NM1) mRNA, complete cds.
ACCESSION AF197940
VERSION AF197940.1 GI:11890405
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1878)
AUTHORS Bolognese,C.P. and McGraw,P.
TITLE The isolation and characterization in yeast of a gene for Arabidopsis S-adenosylmethionine:phospho-ethanolamine N-methyltransferase
JOURNAL Plant Physiol. 124 (4), 1800-1813 (2000)
MEDLINE 20567827
PUBMED 11115895
REFERENCE 2 (bases 1 to 1878)
AUTHORS Bolognese,C.P. and McGraw,P.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1999) Biology, Univ. of Maryland, BC, 1000 Hilltop Circle, Baltimore, MD 21250, USA
FEATURES
source
1..1878
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Landsberg erecta"
/db_xref="taxon:3702"
/chromosome="3"
/map="64921 to 67861"
/tissue_type="stage two leaves"
/clone_lib="Minet M, Dufour ME and Lacroute P., Plant J (1992) 2:417-422"
/dev_stage="complete seedlings"
/note="A. thaliana (L.) Heynh"
1..1878
/gene="NM1"
202..1677
/gene="NM1"
/functions="catalyzes the production of phosphatidylcholine"
/codon_start=1
/product="SAM:phospho-ethanolamine N-methyltransferase"
/protein_id="AAG41121.1"

```

```

/db_xref="GI:11890406"
/translation="MAASYEERDIOKNYIEHSADLTVEAMMLDSASDLDEKPERP
VLSLPPYSGKSVLELACGIRFTQELAQAGELJALDFIDNFKGHSINGHYKVK
PMADVTSPDLKIDGLSDFISNLLAYLSDKEVELLARMVGMIVKGGYIPRESC
PHQSDSKRKNPTHYRBRPYSKVQECOTDADAGNSFELSMIGCKICIGAYVKKON
QNOICWIKVSVSEDRGQRFQVYKSSGILYRERFQGVSTGGLTKEFVR
KMLKPGQVLGVGGGDFYMAKPDVHVVDLDSNMISFALBRALGLSCSVP
EADCTTKHYPDNSPDVYSRDTILHIQDKPALPRTFPMKLPKGVKVLISYCRSPK
SVAFSEYIKRGYDLHDVQAGMLKAGPTDVIARDRTQPMQVLKRELDREKKEK
KATISDFSKEDYDDIVGGWKSLEKRCASDEQKGLPIANK"
361..660
/misc_feature
/gene="NM1"
/note="Region: SAM-binding domain"
361..387
/misc_feature
/gene="NM1"
/note="Region: Motif I"
427..438
/misc_feature
/gene="NM1"
/note="Region: post-I"
544..570
/misc_feature
/gene="NM1"
/note="Region: Motif II"
631..660
/misc_feature
/gene="NM1"
/note="Region: Motif III"
1048..1341
/misc_feature
/gene="NM1"
/note="Region: SAM-binding domain"
1048..1074
/misc_feature
/gene="NM1"
/note="Region: Motif I"
1114..1128
/misc_feature
/gene="NM1"
/note="Region: post-I"
1231..1254
/misc_feature
/gene="NM1"
/note="Region: Motif II"
1312..1341
/misc_feature
/gene="NM1"
/note="Region: Motif III"

ORIGIN
Alignment Scores:
Pred. No.: 1,04e-177 Length: 1878
Score: 2011.00 Matches: 367
Percent Similarity: 88.56% Conservatives: 51
Best Local Similarity: 77.75% Mismatches: 54
Query Match: 79.96% Indels: 0
DB: 8 Gaps: 0

US-10-031-331B-40 (1-473) x AF197940 (1-1878)
QY 1 HisThrValAspLeuThrIleGluAlaMetMetLeuAspSerGlnAlaSerAspLeuAsp 20
Db 256 CATTCGCTGATCTGACTGTTGAGCTATGATGCTGACTGAGAGCTTCTGATCTCGAC 315
QY 21 LysGluGluArgProGluIleLeuSerMetLeuProLeuProLeuGluGlyYsaCysLeu 40
Db 316 AAGGAGAGAGCTCTGAGGTACTCTCTTGTCTCCCTCCATATGAGGCAATCAGTGTG 375
QY 41 GluLeuGlyValAlaGlyIleGlyArgPheThrGlyGluLeuAlaGluYsaAlaGlyGlnVal 60
Db 376 GACCTTGAGACTGGTATGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
QY 61 IleAlaLeuAspPheIleGluSerAlaIleYsaYsaGluValIleGluGlyYsaIleTyr 80
Db 436 ATTGCTCTTGACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495
QY 81 LysAsnValYsaPheMetCysAlaAspValThrSerProThrLeuSerPheProHis 100
Db 496 AAGATGTCAGTTTATGTTGCTGATGTTACATCCCTGACCTCAGATGATGATGATGAT 555
QY 101 SerLeuAspValIlePheSerAsnTrpLeuMetTyrLeuSerAspGluGluValGlu 120

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 04:53:44 ; Search time 128 Seconds
(without alignments)
6945.557 Million cell updates/sec

Title: US-10-031-331b-39
Perfect score: 1602
Sequence: 1 cacaccgtgtattaccat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	149.4	9.3	280	4	US-09-313-294A-6753
2	149	9.3	286	4	US-09-313-294A-7492
3	125.8	7.9	274	4	US-09-313-294A-3443
4	109.2	6.8	260	4	US-09-313-294A-876
5	101	6.3	174	4	US-09-313-294A-5689
6	100.4	6.3	274	4	US-09-313-294A-1188
7	57	3.6	171	4	US-09-313-294A-3925
8	46.6	2.9	7218	1	US-08-232-463-14
9	42.8	2.7	505	4	US-09-621-976-15639
10	42.6	2.7	1106	4	US-09-257-179-19
11	42.4	2.6	5789	3	US-09-242-548-3
12	41.6	2.6	1173	4	US-09-489-039A-2534
13	41.2	2.6	3927	4	US-09-293-238B-1
14	41.2	2.6	3933	3	US-08-199-776-1
15	41.2	2.6	3933	3	US-08-663-731-1
16	41.2	2.6	3933	3	US-08-879-338-1
17	41.2	2.6	3933	5	PCT-US95-02044-1
18	41.2	2.6	1664976	4	US-08-916-421B-1
19	40.6	2.5	1101	4	US-09-252-991A-14068
20	40.6	2.5	1542	4	US-09-252-991A-14202
21	40.4	2.5	1732	1	US-08-480-346-2
22	40.4	2.5	1732	2	US-08-243-541-2
23	40.4	2.5	1732	2	US-08-480-344-2
24	40.2	2.5	2058	2	US-08-749-391-1
25	40.2	2.5	2058	3	US-09-390-200-1
26	39.8	2.5	586	3	US-09-385-982-139
27	39.4	2.5	5852	1	US-07-867-106-2

28 39.2 2.4 946 3 US-08-916-443A-15 Sequence 15, Appl
29 39 2.4 891 4 US-09-134-001C-483 Sequence 483, App
30 39 2.4 1586 4 US-09-178-093B-3 Sequence 3, Appl
31 38.8 2.4 1342 4 US-09-489-847-89 Sequence 89, Appl
32 38.8 2.4 2170 4 US-09-807-258-9 Sequence 9, Appl
33 38.8 2.4 1230025 4 US-09-198-452A-1 Sequence 1, Appl
C 34 38.6 2.4 10968 2 US-08-680-327-2 Sequence 2, Appl
C 35 38.6 2.4 10968 3 US-09-228-246-1 Sequence 1, Appl
36 38.4 2.4 2128 4 US-09-280-116-11 Sequence 11, Appl
37 38.4 2.4 2459 4 US-09-443-795-2 Sequence 2, Appl
38 38.4 2.4 2745 4 US-09-833-381-1410 Sequence 1410, Ap
39 38 2.4 1224 4 US-09-543-681A-262 Sequence 262, App
C 40 37.8 2.4 1164 4 US-09-601-198-65 Sequence 65, Appl
41 37.2 2.3 195 4 US-09-621-976-15314 Sequence 15314, A
42 37.2 2.3 1064 1 US-08-378-588-15 Sequence 15, Appl
43 37.2 2.3 1064 2 US-08-811-094-15 Sequence 15, Appl
44 37.2 2.3 1064 5 PCT-US94-11121-15 Sequence 15, Appl
45 37.2 2.3 1270 1 US-08-378-588-23 Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-313-294A-6753
; Sequence 6753, Application US/09311294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6753
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700352341HI
US-09-313-294A-6753

Query Match 9.3%; Score 149.4; DB 4; Length 280;
Best Local Similarity 71.0%; Pred. No. 5.4e-32;
Matches 198; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 1075 CCAGGAGCTAAAGTTCTTAATCAGTATTACTGCAAGAAAGCTGGTCCACCTCAGTAA 1134
DB 2 CTTGGGGCAGGTCCTTATTCAGTACTCTGCGAGGTCCTGGGAAACCATCAGAGGA 61
QY 1135 TTCGCGCTTACATTAAGCAGAGGGATATGATCTCATGATGTAAAGGAATATGGCAG 1194
DB 62 TTTCAGCGGTACATTAAGCAGAGGGTATGACCTACATCTCTGTGGAGGCTTATGGACAG 121
QY 1195 ATGCTTAAAGATGCTGGATTTGTGATGTTCTTTCGAGGATAGAACTGAGCAGTTCAAT 1254
DB 122 ATGTTGAAGAGTCTGGTTCGTTTCGATGTCATTCGTGAGATCGAATCATGTCCTT 181
QY 1255 CGAGTTCTACGGAAGCACTAGAGATGTTGAGAGGAAGGATGTTTCATTAAGTAT 1314
DB 182 GTGTTTATAGTAGGAGTGTAGCTGAATTTGAAGAGACAGAGCGATTTCTGTCTGAC 241
QY 1315 TTCTTGAGGAGGATTAACATGACATTTGTGGAGGTGG 1353
DB 242 TTCACCCAGAGGACTACGATGATATCTGTTGAACGGTTGG 280

RESULT 2
US-09-313-294A-7492

Sequence 7492, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 7492
LENGTH: 286
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700381824H1
US-09-313-294A-7492

Query Match 9.3%; Score 149; DB 4; Length 286;
Best Local Similarity 70.2%; Pred. No. 7.1e-32;
Matches 200; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 400 GGTACATTTCTTCAGAGATCTGTTTCATCAATCTGGGATCACAACGCAAAAGC 459
DB 2 GGCATATTTCTTGGAGATCATGTTTCCATCTGGAGATTCCTGAGATTCCTGAGAGAGTG 61
QY 460 AATCCACCATCTACCGTGAACCTAGGTTCTACTAAGCCCTTCAAGAGTGTCTATTG 519
DB 62 AACCAACACATCTGAGAACCAAGGTTTATACCAAGGTATTTAAAGAGGCCATTC 121
QY 520 CAGATGATCTGGAACCTCTTATGAGCTCTCCCTACTAGCTGCAAAATGATTGGAGCT 579
DB 122 TTGATCAAGATGGAGTCTGTTGAACTTTCTCTAGTACCTGTAAATGATTGGGCT 181
QY 580 TATGTCAGAACCAAGAACCAAGACAGATTTAGTTGTTGGCAAAAGTTGATTC 639
DB 182 TATGTCAGAACCAAGAACCAAGACAGATTTAGTTGTTGGCAAAAGTTGATTC 639
QY 640 AAGGATGATAAGGGTTCCAGCATTTCTGATCTAGCCAGTAC 684
DB 242 ACAGAGACAGAGATTTCAAAGATTTCTGCAACACGTCGAATAC 286

RESULT 3
US-09-313-294A-3443
Sequence 3443, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 3443
LENGTH: 274
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700611594H1
NAME/KEY: unsure
LOCATION: 52, 202
OTHER INFORMATION: A, J, C, G, or other
US-09-313-294A-3443

Query Match 7.9%; Score 125.8; DB 4; Length 274;

Best Local Similarity 70.6%; Pred. No. 2e-25;
Matches 180; Conservative 0; Mismatches 74; Indels 1; Gaps 1;
QY 655 TTCCAGCGATTTCTGGATATCTAGCCAGTACCAAGTGTAAATAGCAATCTCGCATATGAGCGT 714
DB 5 TTCAAAGATTTCTGGACACAGTGTGATATACCAACAAAGTGGATATTTACGTTACGAGCGT 64
QY 715 GTATTTGCCCTGTTATGTTAGCACTGGAGGATATCAACCAACCAAGAGTTCTGTCA 774
DB 65 GTCTTTGGTGAAGGTTTGTGAGCCTGTTGGAATCGACACCAAGAGATTTGTGGGC 124
QY 775 ATGCTGGACTTGAAGCCCTGCGCAAGCTCTCGATGTTGGTTGTGGAATTTGGTGGAGT 834
DB 125 ATGCTCGATCTTAAACCGGCGCCAGAAAGTACTTGTGATGTCGATGTGGAATTTGGAGCGCGC 184
QY 835 GACTTTTACATGGCGGAGACCTTTGATGT -TGAGGTTGTTGGATTGATCTCTCGGTTAA 893
DB 185 GACTTTTACATGGCTGCNACTATGATGTCTCCATGTTCTGTTATGATCTATCGGTGNA 244
QY 894 TATGATTTCTTTTC 908
DB 245 CATGGTTTCATTTGC 259

RESULT 4
US-09-313-294A-876
Sequence 876, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 876
LENGTH: 260
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6476212 700549985H1
NAME/KEY: unsure
LOCATION: 24, 86
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-876

Query Match 6.8%; Score 109.2; DB 4; Length 260;
Best Local Similarity 68.5%; Pred. No. 7.8e-21;
Matches 178; Conservative 0; Mismatches 80; Indels 2; Gaps 2;
QY 599 ACCAGAACAGATTTAGTTGGTT -GTGCAAAAGTGTGATTTAAGGATGATAAGGGGTTTC 657
DB 1 ATCAAAACAGATATGCTGTTTANTGGGAAAGGTAATAATCAACAGAGACAGAGATTT 60
QY 658 CAGGATTTCTGGA -TACTAGCCAGTACAGTGTAAATAGCAATCTCGCATATGAGCGTCT 716
DB 61 CAAGATTTCTGACACAGCTGCATNCCAAACAGTGGGATTTTTCGTTACGAGCGTGT 120
QY 717 ATTTGCCCTGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGTCAAT 776
DB 121 CTTTGTGGAAGGTTTGTGAGCCTGTTGGAATCGACACCAACCAAGAGATTTGTGGCGCAT 180
QY 777 GCTGGACTTGAAGCCCTGGCAGAGCTCTCGATGTTGGTTGGAATTTGGTGGAGTGA 836
DB 181 GCTCGATCTTAAACCGGCGCCAGAAAGTACTTGTGATGCGATGTGGAATTCGAGCGCGCA 240
QY 837 CTTTACATGGCGAGACT 856
DB 241 CTTTACATGGCGCAACT 260

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pt29pt-F18
US-08-232-463-14

```

US-09-621-976-15639
; Sequence 15639, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15639
; LENGTH: 505
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15639

Query Match          2.7%; Score 42.8; DB 4; Length 505;
Best Local Similarity 17.8%; Pred No. 0.032;
Matches 67; Conservative 144; Mismatches 161; Indels 4; Gaps 1;

QY      741 TGGAGGATATGAACACCAACCAAGAGTTTGTGTCATCTCGACTTGAAGCTGGCCAGAA 800
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      107 YRGWMSKWMKWRKRYMYRMKVCASYSYRRCRKYTGTRGMYGCGKMKCKSTR 166

QY      801 GGTCTCGATGTTGGTGTGTGGAAATGGGGAGGTGACTTTTACATGGCGAGACCTTTGA 860
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      167 YMYTRYTYWMTGACYSKSGMSCKGSRSKYYGSMKYVYGYTWCTSKYSKMSYKSSM 226

QY      861 TGTTCAGGTGTGGATTTGATCTTCGGTAAATATGATTTCCCTTTGCCCTTCGAGGTTTC 920
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      227 CTTTMCYTYWMTCTCTYKKSYYTCKSYTYRSTSKGNGTKSRWSYTWMSKSYTWG 286

QY      921 TATTGGGCTTAATGTGCTGTGTGAGTTTGAGGTAGCAGATGCACCAAGATAAATACCC 980
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      287 CSKKRMKMYMSAGAMYSMMCAKMCAMAGMSMMKCSRAKMYMMAKSYCAKMS 346

QY      981 TGATAA----CTCTTTTGATGTCATCTATAGCGGTGACACCAATTCGCATATTCAGGACA 1036
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      347 CSARSANGRSCTTTTKYKMYKMTTKYCKMYRCMWSKMSCAVCTGYWCYTKTWGTGW 406

QY      1037 AGCTGCGTGTGTTAGATCCTCTCAAAATGCTTGAACCCAGGAGGTAAAGTTCTAATCA 1096
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      407 TAAACTYTGTYKSTCKMNTWMTCTCTWNTWYGTTRATGAGRRGTARRAGACTRAASA 466

QY      1097 GTGATTACTCAGAA 1112
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      467 MTGTTCACTCTTTGAA 482

RESULT 10
US-09-257-179-19
; Sequence 19, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins
; FILE REFERENCE: P2015P1
; CURRENT APPLICATION NUMBER: US/09/257,179
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: PCT/US98/17709
; EARLIER FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/056,270
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,271
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,247
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,073
; EARLIER FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 128

```



```
FILE REFERENCE: L0560/7005/ERP
CURRENT APPLICATION NUMBER: US/09/293,238B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: US 08/879,338
PRIOR FILING DATE: 1997-06-20
PRIOR APPLICATION NUMBER: US 08/663,731
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: US 08/199,776
PRIOR FILING DATE: 1994-02-18
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 3927
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (126)...(3665)
NAME/KEY: sig_peptide
LOCATION: (126)...(179)
NAME/KEY: mat_peptide
LOCATION: (180)...(3662)
US-09-293-238B-1

Query Match 2.68; Score 41.2; DB 4; Length 3927;
Best Local Similarity 67.4%; Pred. No. 0.26; 28; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

QY 1517 TGCATTTCGCAGACTGTAAGATGATTAATCATATTTTATCTTTTAAATTAATCATGATTT 1576
DB 3820 TGCATTTCGCAGAAAGTAAGACTTAGGAACAATTTGGTATTAATAAATTTACACTTTCT 3879

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 3880 TTGCAAAAAAAAAAAAAAAAAAAAAA 3905

RESULT 14
US-08-199-776-1
Sequence 1, Application US/08199776
Patent No. 5594120
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 5594120el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/199,776
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3933 base pairs
STRANDEDNESS: single
TYPE: nucleic acid
```

```
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: mucosal lymphocyte
FEATURE:
NAME/KEY: CDS
LOCATION: 126..3662
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 180..3659
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 126..179
US-08-199-776-1

Query Match 2.68; Score 41.2; DB 1; Length 3933;
Best Local Similarity 67.4%; Pred. No. 0.26; 28; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0;

QY 1517 TGCATTTCGCAGACTGTAAGATGATTAATCATATTTTATCTTTTAAATTAATCATGATTT 1576
DB 3817 TGCATTTCGCAGAAAGTAAGACTTAGGAACAATTTGGTATTAATAAATTTACACTTTCT 3876

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 3877 TTGCAAAAAAAAAAAAAAAAAAAAAA 3902

RESULT 15
US-08-663-731-1
Sequence 1, Application US/08663731
Patent No. 6057423
GENERAL INFORMATION:
APPLICANT: Brenner, Michael B.
APPLICANT: Parker, Christina M.
TITLE OF INVENTION: No. 6057423el integrin alpha subunit
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/663,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,776
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3933 base pairs
STRANDEDNESS: single
TYPE: nucleic acid
TOPOLOGY: linear
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 05:04:23 ; Search time 1100 Seconds

(Without alignments)
7140.755 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602

Sequence: 1 cacacgctgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	834	52.1	1884	13	US-10-425-114-14370
2	822	51.3	1890	13	US-10-425-114-14613
3	743.4	46.4	1897	13	US-10-425-114-24087
4	726.4	45.3	2543	13	US-10-424-599-29759
5	688.8	43.0	1911	17	US-10-437-963-53228
6	676.2	42.2	1962	15	US-10-149-759-81
7	505.2	31.5	1201	13	US-10-425-114-36489
8	446.4	27.9	1267	13	US-10-425-114-2412
9	397.2	24.8	825	13	US-10-425-114-25382
10	388.8	24.3	906	13	US-10-424-599-114653
11	358.2	22.2	594	17	US-10-021-323-14226
12	353.2	22.0	601	17	US-10-021-323-8285
13	329.4	20.6	525	17	US-10-021-323-14357
14	326.6	20.4	694	17	US-10-021-323-17175

15	316.4	19.8	586	17	US-10-021-323-13961
16	303.6	19.0	487	10	US-09-770-961-776
17	290.6	18.1	500	17	US-10-021-323-17176
18	278.8	17.4	467	12	US-09-732-627A-3204
19	255.8	16.0	619	15	US-10-149-759-49
20	238	14.9	557	13	US-10-424-599-97831
21	220.4	13.8	526	13	US-10-424-599-81253
22	209.8	13.1	1087	17	US-10-437-963-7552
23	194.4	12.1	494	10	US-09-770-961-584
24	188	11.7	539	17	US-10-021-323-17177
25	184	11.5	591	17	US-10-021-323-3964
26	177.4	10.8	272	9	US-09-294-093B-1679
27	172	10.7	685	17	US-10-437-963-7551
28	169	10.5	290	9	US-09-294-093B-5022
29	164.2	10.2	477	17	US-10-021-323-6613
30	160.2	10.0	293	9	US-09-294-093B-160
31	156.2	9.8	257	9	US-09-923-876-419
32	152.6	9.5	257	11	US-09-923-876-419
33	152.6	9.5	1416	17	US-10-602-268-18
34	152.6	9.5	1534	17	US-10-602-268-6
35	152.2	9.5	590	17	US-10-437-963-53196
36	150.8	9.4	508	13	US-10-424-599-73778
37	150	9.4	460	13	US-10-424-599-31677
38	144	9.0	261	9	US-09-923-876-984
39	144	9.0	261	11	US-09-923-876-984
40	142	8.9	290	9	US-09-294-093B-2860
41	140	8.7	299	9	US-09-294-093B-6039
42	139.6	8.7	5060	17	US-10-437-963-97616
43	134.8	8.4	274	9	US-09-923-876-3780
44	134.8	8.4	274	11	US-09-923-876-3780
45	132.6	8.3	314	9	US-09-294-093B-4550

ALIGNMENTS

RESULT 1

US-10-425-114-14370
; Sequence 14370, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14370
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB22-074-F3_FLI
US-10-425-114-14370

Query Match 52.1%; Score 834; DB 13; Length 1884;
Best Local Similarity 74.1%; Pred. No. 2.68-219;
Matches 1056; Conservative 0; Mismatches 370; Indels 0; Gaps 0;
Qy 1 CACACGCTGATTAAACATTGAAGCTATGATGTCGATCTCAAGCTTCGATCTTGAC 60
Db 256 CATTCGCTGATCTGACTGTTGAAGCTATGATGTCGATCTTCGAGAGCTTCGATCTGAC 315
Qy 51 CACAGAGAGCTCTGAGATCTTTCATGCTTCGGCTCTTGAGAAATGCTCTTG 120
Db 316 AAGAGAGAGCTCTGAGATCTTTCATGCTTCGGCTCTTGAGAAATGCTCTTG 375

Qy		421	TCTGTGTTCCATCAAACTCTGGGGATACAAAACGAAGCAACTCCACCACCACTACCGTGAA	480
Db		668	TCTGTGTTCCCAACCAATCAGGGGCAGTAAGCGGAAATCCAACCCCACTCACTACCGTGAA	727
Qy		481	CCTAGGTTCTACACTAAGGCCCTTCAAAGAGTGTCAATTTCAGAGATGSGATCTGGAAACTCT	540
Db		728	CCCGGTTCTATTCCNAGGTCCTTTCAGAGGTGCAGACTCGGATGCTCTGGNAATTCA	787
Qy		541	TATGAGCTCTCCCTACTTTAGCTGCAAAATGTATTTGGAGCTTATGTCAGAACAAGAAAAAC	600
Db		788	TTTGAGCTCTCTATGATCGGATGCACTGTGGAGCTTATGTCAAGAACCAAGGAAGAT	847
Qy		601	GAGAACCHAGTATAGTTGGTTGTGGCAAAAAGTTGATCTTAAGSAGTATAAGGGGTTCCAG	660
Db		848	CAGAATCAGATTGTTGGATATGGCAGAGGTCAGCTCAGAAAAATGACAGAGGCTTCCAA	907
Qy		661	CGATTTCTGGGATACTAGCCAGTACAAGTGTAATAGCAATCTGCGATATAGAGCGTGATTT	720
Db		908	CGTTTCTTGGACAATGCCAATACAAATCCAGTGGATCTTAGCTNTGAGCGTGCTCTTT	967
Qy		721	GSCCGTGTGTATGTAGCACCTGAGGATATGAACCAACCAAGAGTTGTGTCAATGCTG	780
Db		968	GCCCCAAGGGTTGTAGCATCTGTGTGACTTTGACAAACCAAGAAATTTGTGAGAGAAATG	1027
Qy		781	GACTTGAAGCTGSCAGAAAGGCTCTGGATGTTGGTTGTGGAATTTGSGAGGTGACTTT	840
Db		1028	AATCTGAACCAAGACAGAAAGTCTTAGATGTTGGGTGGCATTTGGTGGAGGTGACTTC	1087
Qy		841	TACATGGCGGAGACCTTTGATGTTGAGGTGTTGGAATTTGATCTCTCCGTTAANTATNT	900
Db		1088	TACATGGCTGAGAGTTTGAATTCAGTTGTTGGTATCGATCTTTCTGTCAACATGATC	1147
Qy		901	TCCTTTGCCCTTGAGCGTTCTATTGGCTTAAATGTGTGTTGAGTTTGAGGTGACGAT	960
Db		1148	TCTTTCCGATTTGGAACTGTCTATTTGACTCAGCTGCTCGGTTTGAGTTTGAGGTGCTGAT	1207
Qy		961	TGCAACCAAGTAAACTACCCCTGATAACTCTTTTGATGTCACTATATGCCGTGACACCAAT	1020
Db		1208	TGCACCAACMAACACTACCCAGATAAATGGTTTGATGTCATTTACAGCCGTGACACTAT	1267
Qy		1021	CTGCATATTCAAGACAAGCTCGTGTGTTTGTAGATCTTCTACAAATGGTTGAGCCAGGA	1080
Db		1268	CTGCACATCCAGACAAACAGGCTTTGTTGAGCTTTCTTCAAAATGCTTAAACCGGGA	1327
Qy		1081	GGTAAAGTCTATACGTGATTACTGCCAAGAAGCTGGTCCACCTCACTGAAATCGCC	1140
Db		1328	GGTAAAGTCTCATACGCACTACTCTGAGAGCCCAAACTCCATCTGCTGAGTTTTCA	1387
Qy		1141	GCTTACATTAAAGCAGAGGGGATATGATCTCCATGATGTAAAGGAATATGGCCAGATGTT	1200
Db		1388	GAGTACATCAAAACAGAGAGATATGATCTCCATGACGTTCAAGCTTATGGCAAGTCTTA	1447
Qy		1201	AAAGATCTCGAATTTGTGTGATGTTCTTCGAGGATAGAACTGAGCAGTTCAITTCGAGTT	1260
Db		1448	AAAGCGCTGGCTTCTAGTATGTATGGCAGAGGACCGTACTGATCAGTTTATGCAAGTC	1507
Qy		1261	CTACGGAAGCACTAGAGACTGTTTGAGAGGAAAAAGGATGTGTTCAATTAGTGAATTTCTCT	1320
Db		1508	CTGAAACGTGAATTAGACAGGGTGGAGAAAGAAAAAGGAAAAATTCATCTCCGACTTCCC	1567
Qy		1321	GAGGAGATTAACAATGACATTTGTGAGGTTCGAATGATAGTTTCGGAGGACTGCGCAG	1380
Db		1568	AAAGAGGNTTACGNATGNCATTTGTGNGNTGATAGTCAAAAGCTGGAGAGGTGTGCATG	1627
Qy		1381	GGTGAGCAACGATGGGGTCTGTTCGTTGCCAAGAAAGTCAAGAA	1426
Db		1628	GATGAGCAGAAATGGGGACTTTTTCATCGCCAAACAGAAATTAAGCA	1673

RESULT 3

Record 3
 US-10-425-114-24087
 ; Sequence 24087, Application US/10425114
 Publication No. US20040034888A1
 Db
 661 CGATTTCTGGATCTACGCCAGTACAGCTGTAATAGCAATCTCGCATATCAGCGTGTAATT 720
 900 AGTTTTTTGGCAAAATGTCAGTACAAACCACTCGGAATATCTAGCTATGAACGCTACTCTT 959

Db 1613 ACACCATGCTGCAGCTCAAGATATAACCAACATATTTACAGATCAATTTTACAAGTGGTTGA 1672
Qy 1073 AGCAGGAGGTAAGATTTCTAATCAGTGAATTTACTGCAAGAAAGCTGGTCCACCTCACCTG 1132
Db 1673 AGCCTGGAGGAAATTTCTAATCACAGATTTACTGCAAAAGTCTGGAAGTCACTTTAG 1732
Qy 1133 AATTCGCGCTTACATTTAAGCAGAGGGGATATGATCTCATGATGTAAGGAATATGGC 1192
Db 1733 AATTTGCTGAGTACATATAAAGGAGGATATTTATCTCATGATTAAGCGTATAGGC 1792
Qy 1193 AGATGCTTAAAGATGCTGGATTTGTTGATGTTCTTGCCGAGATAGAACTGAGCAGTTCA 1252
Db 1793 AGATGCTTGAAGATGCGGATTTGATGATGCTCATTTGCCGAGATCGAACTGATCAGTTG 1852
Qy 1253 TTGAGTCTTACGGAAGGACTAGAGCTGTTGAGGAAGAAAGATGTTGTTCAATGATG 1312
Db 1853 TGACACGCTACAGCAGAGTTAATGCTTGGAGAACGAAGACCATTTTATTTGTTG 1912
Qy 1313 ATTTCTTGAGGAGGATTAATGATGATTTGTTGGAGTTTGGAAATGATTAAGTTGCGGAGGA 1372
Db 1913 ACTTCAGCGAGGAAGACTACATGAGATTTGTCGAAAGATGGAAGCAAGCAGACCGAGG 1972
Qy 1373 CTGCAAGGCTGAGCAACGATGGGCTGTTGCTTCCCAAGAAAGTGA 1422
Db 1973 GTGCACTAGAGCAGATGCGGGCTTGTTCATTTGCCAAGAAAATTTGA 2022

RESULT 5

US-10-437-963-53228
; Sequence 53228, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yibua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boucharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 53228
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5544C.1
US-10-437-963-53228

Query Match 43.08; Score 688.8; DB 17; Length 1911;
Best Local Similarity 69.48; Pred. No. 3.7e-179;
Matches 967; Conservative 0; Mismatches 397; Indels 30; Gaps 1;
Qy 79 ATTCCTTCAATGCTTCGCTCTTGAAGGAAATGCTCTTGGAACTTGGGGCTGGTAT 138
Db 223 ATATTTGCTTTACTTCTCTCTTACGAAAGGAAATCAGTACTGGAACCTTGGTCTGGAATA 282
Qy 139 GTCCTTTTACTGCTGAATTTGGCTGAGAAAGCTGGCAGGTTATTTGCTCTGATTTTCATT 198
Db 283 GTCCTTCTACTGGAAGTATGTAAGAACGCTGGGATGTTCTTGAATGATTTTCATT 342
Qy 139 GAGAGTCTTCAAGAGAGATGAGTAAATCAATGGGCACTACAAAATGTCAGATTTATG 258
Db 343 GAAAGTGTGATTAAGAGATGAAGCATAAAGGCTCACCAAGATGTCATCTCTTTATG 402
Qy 259 TGTGCTGATGATCTTCTCCACTCTCAGTTTCCACCACTTCATTTGATGATGATATTC 318
Db 403 TGTGGGATGTCACATGCTCCAGACCTGATGATGAGGATTAATCTCCATTTGATCTGATTT 462

Qy 319 TCCAAATGGTTACTCATGATATCTTTCTGATGAAGAGGTGGAATAATTTGGTTGAAGAATG 378
Db 463 TCAAACTGGTTACTGATGATCTTTCTGAGCAGAGAGGTTGAGAAGCTAGTAAAGGAATG 522
Qy 379 TTGAAATGGTTGAAGCAGAGGGGTTACATTTTCTTCAAGAGAACTTTGTTTCCATCAATCT 438
Db 523 GTAAAGTGGCTAAAGGTTGGCGCTATATCTTCTTTAGGGAATCTTGTTCATCAGTCT 582
Qy 439 GGGGATCAAAACCAAAAGCAATCCACCACTACCGTGAACCTAGTTCCTACACTAAG 498
Db 583 GGAGTTCAAAAGGAAAGTGAATCTTACATATCCGGGAGCCCAAGGTTTACACTAAG 642
Qy 499 GCCTTCAAGAGTGTCTATTTGGAAGATGGAATCTGGAACCTTATATGAGCTCTCCCTACTT 558
Db 643 GTGTTTAAAGAGTGTCAAGCTCTTGATCAAGATGGGATTTCCCTTTGAATCTCTCTGTACTT 702
Qy 559 AGCTGCAATATGTTGGAGCTTATGTGAGAAACAAAGAAACCAAGAACAGATAGTTGG 618
Db 703 ACTTGAAGTGTGTTGGAGCTTACGTGAAAGCAAGAAATCAAAACCAAGATATGG 762
Qy 619 TTGTGGCAAAAGCTTGATTTCTAAGGATGATAAGGGGTTCCAGCGATTTCTGATACTAGC 678
Db 763 CTATGGCAAAAGGTTGATTTCAAGAGATCGGGGTTTCAAGATTTTGGCAATGTTG 822
Qy 679 CAGTACAAAGTGTATAGCATTTGCGATATGAGCGTATTTGGCCCTGGTTATGTTAGC 738
Db 823 CAGTACAAAGCCAGTGGATATTTACGCTATGAACGCATCTTTGGAGAGGGCTTTGTGAGC 882
Qy 739 ACTGGAGGATAT-----GAAACCAACCAAGAGATTT 768
Db 883 ACTGTGGAATTTGATGCTTATTTTATCTTCTGTTCAACCAAGAACTTACAAAGATTT 942
Qy 769 GTGTCAATGCTGCACTTTGAAGCTGCGCAGAGAGTCTCGGATGTTGGTTGGAATGGT 828
Db 943 GTGGACAGGCTGGATCTCAAACTGCGCCAGAACGTTCTTGATGTTGGATGGAATGGG 1002
Qy 829 GGAGTGACTTTTACATGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCC 888
Db 1003 GGCGTGATTTTATATGCTGACAAAGTATGATGTTCTATGTTGTTGTTGATTTGCTTCG 1062
Qy 889 GTTAATATGATTTCTTTGCTTGGAGCTCTTATGCGCTTAATGCGCTTAATGCTGGATTT 948
Db 1063 ATAAACATGTTTCTTTGCACTTGAGCGTCTATTTGGCGTAAAGTCTCAGTTGAGTTT 1122
Qy 949 GAGTACGAGATTCACCAAGATAAACTACCTGATAACTCTTTTGAATGTCATCTATAGC 1008
Db 1123 GAAGTCTGATTTGCAAAAAGACATACCCAGACACACGTTTGAAGTCTATCTACAGT 1182
Qy 1009 CGTACACCATTTCTGATATTCAGGACAAAGCTTGGTGTGTTAGATCCTTCTCAAAATGG 1068
Db 1183 CGTGATCTATCTTCAATACAGATTAACCTCACTATTTAAAGTTTCTTCAAGTGG 1242
Qy 1069 TTGAGCCAGAGTAAAGTTCTTAATCAGTGAATTAATGCAAGAAAGCTGGTCCACCTCA 1128
Db 1243 CTCAAACTGGGGTAAAGTCTTAATTAGTGAATTAATGAGTGGCTTGGGAAACCTTCA 1302
Qy 1129 CTTGAATTCGCGCTTACATTAAGCAGAGGAGATGATCTCCATGATGTAAGAAATAT 1188
Db 1303 GAAGATTCGAGCTTACATTAAGCAAGGGTTATGACCTTTCACGAGCTCAGGGCTAC 1362
Qy 1189 GGGCAGATGCTTAAAGATGCTGATTTGTTGATGTTCTTGGGAGGATGAGAACTGAGCAG 1248
Db 1363 GGAAGATGCTTGAAGATGCTGTTTCCATGATGATGTTGCTGAAGACCGCAGGATCAG 1422
Qy 1249 TTCAATCGAGTTCTACGGAAGAACTAGAGACTGTTGAGAAAGAAAGAGATGTTCTTCAAT 1308
Db 1423 TTCTCGATGTTCTAGAGAGGAGCTTGTCTAAAGTTGAAAGAACAAACGAGTCTGTC 1482
Qy 1309 AGTGATTTCTCAGAGAGGATTAAGATGATGTTGAGGTTGGAATGATAGTGGG 1368
Db 1483 TCTGATTTGAGCCAGAGAGACTAGAGCCATGTAATGATGAGGCAAAACTTCAA 1542

QY 1369 AGGACTGCCAAGGCTGAGCAAGATGGGGTCTGTGTTGCGAAGAGAGTGAAGATC 1428
DB 1543 AGGAGTTCTGCTGGTGAGCAGAGTGGGGCTGTTTCATCGCGACCAAGTGAATCAAG 1602
QY 1429 AGTTGCGGCACTGG 1442
DB 1603 ATGCGCAGCTGGG 1616

RESULT 6

US-10-149-759-81
; Sequence 81, Application US/10149759
; Publication No. US20030157592A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Reinhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins
; TITLE OF INVENTION: involved in the synthesis of tocopherols and
; TITLE OF INVENTION: carotenoids.
; FILE REFERENCE: BASF/NAE 1333/99 PCT/US
; CURRENT APPLICATION NUMBER: US/10149, 759
; PRIORITY FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/EP/00/12698
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Wordperfect version 6.1
; SEQ ID NO 81
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Physcomitrella patens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (367)..(1842)
; OTHER INFORMATION: 78_ppprotol_092_e12-260rev
US-10-149-759-81

Query Match 42.2%; Score 676.2; DB 15; Length 1962;
Best Local Similarity 68.1%; Pred. No. 1.2e-175;
Matches 974; Conservative 0; Mismatches 448; Indels 9; Gaps 2;

QY 1 CACACGCTTGATTAACCATTAAGCTATGATGCTGATCTCAAGCTCTCTGATCTTGAC 60
DB 415 CATTCTGTGGAGCCTAGGCTGAGCAATGATGCTTGAATGCGAGCCTCCAACTCGAT 474
QY 61 AAGAAGAAAGCTCTGAGATCTCTTCAATGCTTCCGCTCTTGAAGGAAATGCCCTCTTG 120
DB 475 AAGAAGAAAGCAACCGAGATTTGCTGCTGTGCGCCCATATGAAACAAAGATGTCATG 534
QY 121 GAATCTGGGGCTGGTATGCTGCTGTTTACTGCTGATGCTGCAAGCATGCAAGTCTATGTG 180
DB 535 GAGCTCGGAGCAGGCATCGTGGTGTACTGCTGCTGCAAGCATGCAAGTCTATGTG 594
QY 181 ATGCTCTGATTTCAATGAGAGTCTATCAAGAGAGTGAATCAATGCGCACTAC 240
DB 595 CTGCGCATGATTTCAATGAGATCTCATCAAGAGAGTGAATCAATGCGCACTAC 554
QY 241 AAAAAATGCAAGTTTATGCTGCTGATGCTCTCTCCACTCTCAAGTTCCCAACACAT 300
DB 655 AACAAATGCAATTTCAATGCTGCTGATGCTCTCTCCACTCTCAAGTTCCCAACACAT 714
QY 301 TCATTTGATGATGATTTCTCAATTTGGTATCTCATGATCTCTCTCAAGAGGTTGAA 360
DB 715 TCTGCGGATCTGCTGTTTCAATTTGCTTCTCATGATCTCTCTCAAGAGGTTTAA 774
QY 361 AATTTGTTGAAGATGTTGAATGCTTGAAGGAGGTTTCAATTTCTCTCAAGAA 420
DB 775 GGCCTAGATCAGCGGTATGAGGTGCTCAGGCTGAGATCATTTCTCTCAAGAA 834
QY 421 TCTTTTCCATCAATCTGGGATCACAACGCAAGCAATCCCAACCACTACCGTGA 480

DB 835 TCCTGCTTCCACAGCTCAGGAGATCAAGCGAAAGAAATCTCTACTACTACCGTGA 894
QY 481 CCTAGGTTCTACACTAAGCGCTTCAAGAGTGTCTATTGTCAGAGTGAATCTGGAATCTCT 540
DB 895 CCCAACGAGTACACGAACATCTTCCAGCAGGCTTACATCGAAGAGG---ATGGGTCCTAT 951
QY 541 TATGAGCTCTCCCTACTTATAGCTGCAAAATGATTTGAGCTTATGTCAGAAACCAAGAAAC 600
DB 952 TTCAAGTTTGAATGGTCCGATGCAAAATGTTGCGCATATACGTCGCGAATTAAGAAAT 1011
QY 601 CAGAACCAAGATTAGTTGGTTGGGCAAAAGTTGATTTCTAA-----GGATGATAAGGG 654
DB 1012 CAAACCCAGGTGTGTGGTTATGAGAGAAAGTTTCACTCGGATGGAACCTGAGAGCGAGTGT 1071
QY 655 TTCCAGCGATTTCTGGATACTAGCAGTACAAAGTGTAAATAGCAATCTCCGATATAGCGGT 714
DB 1072 TTCCAGAAAGTTTGGACACCCCAACAGTACACGTCAACTGGAATCTCTGGTTACAGCGT 1131
QY 715 GTATTGGCCCTGTTATGTTAGCACTGAGAGATGAAACCAACCAAGAGTTTGTGTCA 774
DB 1132 ATTTTGGAGAGGATTTGTAGCAGGTTGGAATCGAAACCAACCAAGAGTTTGTAGT 1191
QY 775 ATGCTGGAATTGAAGCTCGCCAGAAAGTCTCTGGATTTGGTTGTGGAATTTGGTGGAGT 834
DB 1192 ATGCTGGAATTGAAGCTCGCCAGAGCAGCGTCTCTGAGCTTGGATGTGGGATCGAGGTGT 1251
QY 835 GACTTTTACATGCGGAGAGCTTTGATGTTGAGGTTTGGATTTGATCTCTCCGTTAAT 894
DB 1252 GATTTTACATGCGGAGAGATGATGATGCTGAATTTGTCGATGCTGCTCTTAAT 1311
QY 895 ATGATTTCTTTCCTTGGCTTGGAGCTTCTAATGGCTTAAATGGCTTGGTTGAGTTGAGTA 954
DB 1312 ATGATTTCTTTCCTTGGCTTGGAGCTTCTAATGGCTTAAATGGCTTGGTTGAGTA 1371
QY 955 GCAGATTGCAACCAAGATAAATCTACCTGATAATCTTTTGTGATGTCATCTATAGCGGTGAC 1014
DB 1372 GGGGATTTGCAACCAAGATTAATACCTCTACGATCTTTTGTGATGTCATCTACAGTCTGTAT 1431
QY 1015 ACCATCTGCTATATTAGGCAAGCTCGGTTGTTAGATCTCTTACAAATGGTTGAG 1074
DB 1432 ACCATCTGCTATATTAGGCAAGCTCGGTTGTTAGATCTCTTACAAATGGTTGAG 1491
QY 1075 CAGGAGTAAAGTTCTTAATCAGTGAATCTGCAAGAAAGCTGCTCCACCTCACTGAA 1134
DB 1492 CTTGAGGCTCGGCTGCTTATCAGTGAATCTGATAGAGCTCCACAACTCCGTCGCGGAG 1551
QY 1135 TTGCGCGCTTACATTAAGCAGAGGAGATATGATCTCCATGATGTAAGGAAATATGCGAG 1194
DB 1552 TTGCTGCTATATTTAGCAGAGGAGTATGATCTCATAGGCTTCAAGATGAGAGAG 1611
QY 1195 ATGCTTAAAGATGCTGAATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCAAT 1254
DB 1612 ATGCTGGAAGATGCGGTTTGTGGAAGTGTGCGAGAGGACCGCAAGATCAGTTCAAT 1671
QY 1255 CAGATTCTACCGAAGGAATAGAGACTGTTGAGAGGAAAGAGATGTTTCAATTAAGTGAAT 1314
DB 1672 GAAGTTTACAGAGGAGCTAGCCACCACTGAAGCAGGCTGTCGACAGTTTCAACAGAT 1731
QY 1315 TTCTGAGGAGGATTAACATGATGATTTGAGGAGTTGGAATGATGATGTCGAGGAGT 1374
DB 1732 TTCTCGAGGAGGATTAACATGATGATTTGAGGAGTTGGAATGATGATGTCGAGGAGT 1791
QY 1375 GCCAAGGAGTGAAGCAAGATGGGCTCTGTGCTTGGTGGCAAGAGAGTGAAGA 1425
DB 1792 TCGAATGCAAGAGAGTGGGAGCTCTCTCATAGCTTACAGGCAATATGA 1842

RESULT 7

US-10-425-114-36489
; Sequence 36489, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 36489
LENGTH: 1201
TYPE: DNA
ORGANISM: Zea mays subsp. mexicana
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE108B03_FLI
US-10-425-114-36489

Query Match 31.5%; Score 505.2; DB 13; Length 1201;
Best Local Similarity 71.0%; Pred. No. 1.6e-128;
Matches 669; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 475 CGTGAACCTAGGTTCTACACTAAGGCTTCAAGAGTGTCTATTGCAAGATGGATCTGGA 534
DB 1 CGAGAACCAAGTTTATACCAAGGTATTTAAGAGGSCCATTCATTGTATCAAGATGGA 60
QY 535 AACTCTTATAGCTCTCCCTACTAGCTCAATGTATTTGGAGCTTATGTGCAAGAACAG 594
DB 61 GGTTCGTTTGAACCTTCTCTAGTGACCTTAAATGTATTTGGGCTTATGTCAAAAACAG 120
QY 595 AAAAAACAGAACAGATTAAGTTGTTGTCGCAAAAAGTTGATTTCAAGCATGATAAGGG 654
DB 121 AAGAAATCAAAACAGATATGCTGTTATCGGAAAGAGTAAATCAACAGAGACAGAT 180
QY 655 TTCCAGCGATTTCTGATACCTAGCTACAGTGTATAGCAATCTCTCGATATACCGT 714
DB 181 TTTCAGAGTTCTGCAACAGCTGCANTACAAACAGTGGGATATACGTTATGAGCGT 240
QY 715 GTATTTGGCCCTGTTATGTTAGCACTGGAGGATATGAACCAACAAAGAGTTTGTGCA 774
DB 241 GTCTTTGCTGAAGTTTGTGAGCACTGCTGGAATCGAGACTACAAAGGAATTTGTGGC 300
QY 775 ATGCTGAGCTTGAAGCTCGCCAGAGGCTCTGGAAGTGTGTTGGAAATTTGGTGGAGT 834
DB 301 ATGCTGAGCTTAAACCGGSCAGAAAGTACTTGATGCGATGTGGAATTTGGAGCGC 360
QY 835 GACTTTACATGGGGAGCTTTGATGTTGAGTGTGTTGGAATTTGATCTCTCGTTAAT 894
DB 361 GACTTTACATGGCTGCAACTATGATGTCCTGATGTCCTGTTGATTTGATCTTTGCGTGA 420
QY 895 ATGATTTCTTTCCTTGGAGCTTCTATTTGGGCTTAAATGTCCTGTTGAGTTTGAAGTA 954
DB 421 ATGCTTTCTATTTGCAATTTGAACCTGCAATTTGACGCAAGTGTCTCTGTTGAATTTGA 480
QY 955 GCAGATTGCAACAGATTAACCTGATTAATCTTTTGAATGCTATAGCGGTGAC 1014
DB 481 GCTGTTGCAACCAAGAGTTTACCAGAAATAGTTTGAAGTCACTCTACAGCCGTGAC 540
QY 1015 ACCATTCTGATATTCAGGCAAGCTCGCTGTTGTTTAGATCTCTCAAAATGTTGAAG 1074
DB 541 ACCATCTCTCATAACAGCAAGCGCTCTGTTTGAAGCTTCTCTCAATGCTTAAG 600
QY 1075 CCAGAGGCTAAGTTTCTAACTAGTATTCGCAAGAGCTGTCACCTCTCACTGTA 1134
DB 601 CCGGCGGCAAGCTCTATACGCACTACTGTAGATCTCTGGAATCCCTGGAATCAGAGAA 660
QY 1135 TTGCGCGCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAAGGAATATGGCAG 1194
DB 661 TTGCTGCTGATCATTAAGCAGAGAGGCTATGACCTTACGAGCTGAAGCTTATGACAG 720
QY 1195 ATGCTTAAAGATGCTGATTTGTTGATGTTCTTTCGCGAGATAGAACTGACGATTCATT 1254

RESULT 8

US-10-425-114-2412
Sequence 2412, Application US/10425114
Publication No. US20040034888A1

GENERAL INFORMATION:
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 2412
LENGTH: 1267
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700211781_FLI
US-10-425-114-2412

Query Match 27.9%; Score 446.4; DB 13; Length 1267;
Best Local Similarity 72.0%; Pred. No. 2.9e-112;
Matches 582; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 614 GTTGGTTGCGCAAAAGTTGATTTCAAGATCATAGGGGTTCCAGCGATTTCTGATTA 673
DB 191 GTTGGCTATGAAAAGTAACCTCATCAGAGATGGGGATTTCAAGTTTCTTGACA 250
QY 674 CTAGCCAGTCAAAAGTGTAAATAGCAATTCGCGATATAGCGGTGATTTGGCCCTGTTATG 733
DB 251 ATGTGCAGTCAAAAGCCACTGGAATACTACGCTATGACGATCTTTGGAGATGGCTAG 310
QY 734 TTAGCACTGCGAGCATATGAACCCACCAAGAGTTTGTGTCAATGCTGAGCTTGAAGCTG 793
DB 311 TGAGTACTGCTGAGCTGAGACTACAAAGAAATTTGAGAGAACTGAATCTTAAGCTG 370
QY 794 GCCAAGAGTCTCGGATGTTGGTGTGGAATTTGGTGGAGTGACATTTTACATGCGGAGA 853
DB 371 GGCAGAGGTGCTTGATGTTGGATGTGGAATTTGGGAGGAGTGACATTTTATATGCTGAGA 430
QY 854 CTTTGAATGTTGAGGTTGTTGGAATTTGATCTCTCCGTTAAATATGATTTCTTTGCCCTTG 913
DB 431 AGTATGCTACACATGTTGTTGTTATTTGACCTTTTCCATTAACATGATATGTTGCCCTTG 490
QY 914 AGCTTCTATTTGGCTTAATGTTGCTGTTGAGTTTGAAGTGTAGGAGTGTGCAAGATTA 973
DB 491 AGCTTCTATTTGGGTTGAAGTGTGCTTGAAGTGTGCTGATTTGACCAAGAA 550
QY 974 ACTACCTGATTAATCTTTTGTGATGTCATCTATAGCGGTGACCACTTCTGCAATTTCCAG 1033
DB 551 CATACCCAGACCAATGTTGATGTCATCTACAGTGGTGAACACTATCTCTCATATACAG 610
QY 1034 ACAAGCTGCTGTTGTTAGTCTCTTACAAATGTTGAGCGAGGATTAAGTTCTTA 1093

Db 611 ATAAACCTCTCTGTTTAAAGTTTCTTCAATGGCTGAACCTGGGGAAAGGTTCTAA 670
QY 1094 TCAGTGATTACTGCAAGAAAGCTGGTCCACCTCACCCTCACTCAATTCGCGCTTACATTAAGC 1153
Db 671 TCAGTGATTACTGCAAGAGTCTCTGAAACCATCAGAGGTTTGAACATACATTAAGC 730
QY 1154 AGAGGGGATGATCTCATGATGTAAGGATATGCGGAGATGCTTAAAGATGCTGGAT 1213
Db 731 AGAGGGGATGATCTCATGATGTAAGGATATGCGGAGATGCTTAAAGATGCTGGAT 790
QY 1214 TTGTTGATGTTCTTTCGCGAGGATAGAACTGAGCAGTTCATTCGAGTTCCTACGGAAGAAC 1273
Db 791 TCAGTCAATGCTTTCGGAAGCCGAACTGACCAAGTCTCTCAGTCTTTCACGAAGGAGC 850
QY 1274 TAGAGACTGTTGCAAGGAAAGGATGCTTTCATGATGATTTCTCTGAGGAGGATTAACA 1333
Db 851 TAGACAAATTTGAGAGAACCAAGATGATTTCTCTGCTGATGTTTCCCGAGGAGGATTAAG 910
QY 1334 ATGACATGTTTGGAGGTTGGAATGATTAAGTTGCGAGGAGCTGCGCAAGGATGAGCAAGAT 1393
Db 911 AGGATATGCTGAATGGAATGGAAGCAAACTGCAAGAGGAGCTCTGCTGAGGAGGAGGT 970
QY 1394 GGGGCTCTGTTGTTGCCAAGAAAGGATG 1421
Db 971 GGGGCTCTGTTGTTGCCAAGCAAGGATG 998

RESULT 9

US-10-425-114-25382
; Sequence 25382, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25382
; LENGTH: 825
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3829-031-F4_FLI

Query Match 24.8%; Score 397.2; DB 13; Length 825;
Best Local Similarity 69.4%; Pred. No. 8.8e-99;
Matches 568; Conservative 0; Mismatches 248; Indels 2; Gaps 2;
QY 762 AGAGTTTGTGCAATGCTGAGCTTGAAGCTTGGCAGAGGCTCTGGATGTTGGTTGG 821
Db 1 AGAATTTGGCGAAGTTGGATCTAAAGCTTGGCCAAAGGTTCTAGATGTTGGCTGG 60
QY 822 AATGTTGGAGTGATCTTATCATGCGGAGACCTTTGATGTTGAGTGTGTTGGAATTTGA 881
Db 61 CATAGTGGAGGTGACTTTTATATGCGCAAGAAATTTGAGTCTATGTTTGGGCATCGA 120
QY 882 TCTCTCCGTTAATGATTTCTTTGCTTGGCTTGGAGCTTCTATTTGGGCTTAAATGCTGT 941
Db 121 TCTGTCATTAATGATTAATCTTTGCTTCTTGAAGGCAATGAGTCTGATGTTCTCTGT 180
QY 942 TGAGTTGAGGTTGAGGATGCAAGATTAACCTACCTGATACTCTTTTGAATGTCAT 1001
Db 181 GGAATTTGAAGTTGCTGATGCTGATGCTACAAAGACTTATCCAGATAACAGCTTTGAT 240

QY 1002 CTATAGCGGTGACACCACTTCTGCATATTCAGGACAGCCCTGGTGTGTTAGATCCTCTTA 1061
Db 241 CTACAGCGGTGACACCACTTCTGCATATTCAGGACAGCCCTGGTGTGTTAGATATTTCTA 300
QY 1062 CAATATGTTGAGCCAGGAGTAAAGTCTTAATCAGTGAATTAATCTCAAGAGAGCTGGTCC 1121
Db 301 CAATATGTTGAGCCAGGAGTAAAGTCTTAATCAGTGAATTAATCTCAAGAGAGCTGGTCC 360
QY 1122 ACCCTCACTGAATTCGCCCTTACATTAAGCAGAGGAGGATATGATCTCCATGATGTA 1181
Db 361 TTATCGCCCGCAATTTGCAAGTACATCAAGCAGAGAGGCTATGATCTCCATGATGTA 420
QY 1182 GGAAATATGGGAGAGATGCTTAAAGATGCTGATTTGTTGATGTTCTTGGCGAGGATAGA 1241
Db 421 AGCTATATGGCCAGATGCTTAAAGATGCTGATTTGATCAGGTCAATTCAGAGGATCGAA 480
QY 1242 TGAGCAGTTCATTCGAGTCTTACGGAAGGAACTGAGAGCTGTTGAGAGGAAAGGATGT 1301
Db 481 CGATCAGTTCATTCGAGTCTTACGAGGCTGAATTAGACAAAGTTGAGAGGAAAGGATG 540
QY 1302 GTTCATATGATGATTTCTCTGAGGAGGATTAATGATGATTTGAGGTTGGAATGATA 1361
Db 541 CTTCATCAGGAGTCTTCTCAGGAGGATTAATGATGATTTGAGGTTGGAATGATA 600
QY 1362 GTTGGGAGGAGTCTGCAAGGCTGAGCAACGATGGGCTGTTGCTGTTGCCAGAGAAAGTG 1421
Db 601 GCTGATTAGGACCAAGATTTGGGAGGAGGAGATGGGCTCTTTTCATTTGCCACAAACTA 660
QY 1422 AAGAT-CAGTTGGGAGTCTGAGTCTGATGATTTCTAGTATTAATCTTCAATGTTTCA 1480
Db 661 AGTCTGCTTCTCTCTAATGTTTACATTTCTCATTTCTTAATTAATGATTCATAGTA 720
QY 1481 TGTAATGATCTTCTACATGTAATAATTTGCCAATGATTTGCAATTTGCGAGAGTGTAG 1540
Db 721 AG-AGTGTACTTAGCATAGATGTTTCTACTGATGCTTGTGCTCTGATGTTATGAAC 779
QY 1541 TTAATCATATTTTATCTTAAATTAATCATGGAATTTAT 1578
Db 780 TTAACATCTTTTGGCTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 817

RESULT 10

US-10-424-599-114653
; Sequence 114653, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 114653
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74541C.1

Query Match 24.3%; Score 388.8; DB 13; Length 906;
Best Local Similarity 71.2%; Pred. No. 2e-96;
Matches 513; Conservative 0; Mismatches 207; Indels 0; Gaps 0;
QY 709 GAGCTGATTTGGCCCTGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 768
Db 1 GAGCTGATTTGGCCCTGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 769 GTGTCAATGCTGAGCTTGAAGCCTGGCAGAGGCTCTGATGATGTTGTTGTTGTTGTT 828

```

Db 61 GTGGCAAGCTGGGACATAAACCTGGCCAGAAAGTACTGGAGTGTGGTACTGGG 120
Qy 829 CGAGGTGACTTTTACATGGCGGAGACCTTTGATGTTGGAGTGTGGATTTGATCTCTCC 888
Db 121 GAGGTGACATTTACATGGCAGAAATTTGATGTTGGAGTGTGGATTTGATCTCTCC 180
Qy 889 GTTAATATGATTTCTTGGCCCTTGGAGCTTCTAATTTGGGCTTAATTTGCTGTGAGTTT 948
Db 181 ATAAACATATTTCTTGGCCATTTGAACGCTGCTAATTTGGAATCTCAATGCTGTGAGATTT 240
Qy 949 GAGGTAGCAGATTGACCAAGATATAACTACCTGATATACTTTTGAATGCTATATAGC 1008
Db 241 GAGTGTGGCGATTGCACTTAAACAAACATTTCTCTGTAATACATTTGATGTAATCTATTC 300
Qy 1009 CGTGACACATCTGCAATTTACGACAGAGCTGCTGTTGTTAGATCTTCTACAAATGG 1068
Db 301 CGTGACACATCTGCAATTTACGACAGAGCTGCTGTTGTTAGATCTTCTACAAATGG 360
Qy 1069 TTGAAGCCAGGAGGTAAAGTTCTAATCAGTGAATTAAGTCTGCAAGAAAGCTGCTCA 1128
Db 361 TTGAAGCCAGGAGGTAAAGTTCTAATCAGTGAATTAAGTCTGCAAGAAAGCTGCTCA 420
Qy 1129 CTTCAATTCGCGCTTACATTAAGCAGAGGAGATATGATCTCCATGATTAAGGAATAT 1188
Db 421 TTAGGATATCTGCAATTAAGCAGAGGAGATATGATCTCCATGATTAAGGAATAT 480
Qy 1189 GGGCAGATGCTTAAGATGCTGGATTTGTTGATGTTCTTCCGAGGATAGACTGAGCAG 1248
Db 481 TGTGGATGCTCGAATGCTGGATTTGTTGATGTTCTTCCGAGGATAGACTGAGCAG 540
Qy 1249 TTTCAATTCGAGTTCTACGGAAGGAACTAGAGACTGTTGAGAGGAAAGATGTTCAAT 1308
Db 541 TTTCAATTCGAGTTCTACGGAAGGAACTAGAGACTGTTGAGAGGAAAGATGTTCAAT 600
Qy 1309 AGTGAATTTCTGAGGAGATTAAGTGAATGATGTTGAGGATGATGATGATGATGATG 1368
Db 601 GATGACTTCTGAGGAGATTAAGTGAATGATGTTGAGGATGATGATGATGATGATGATG 660
Qy 1369 AGGACTGCAAGGCTGAGCAGATGAGTGGGCTGTTCTGTCGCAAGAAAGATGAGAAATC 1428
Db 661 CGGGGGCGGAGATGTTGAGCAAAATATGAGGCTTGTTCATGTCGCAAGAAAGATGATC 720

```

RESULT 11

```

US-10-021-323-14225
; Sequence 14225, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR FILING DATE: 2001-12-12
; PRIOR FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 14225
; LENGTH: 599

```

```

; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(599)
; OTHER INFORMATION: unsure at all n locations

```

```

US-10-021-323-14225
; Query Match 22.2%; Score 355.2; DB 17; Length 599;

```

```

Best Local Similarity 75.7%; Pred. No. 2.9e-87;
Matches 452; Conservative 0; Mismatches 144; Indels 1; Gaps 1;
Qy 761 AAGATTTGTGTCAATGCTGGACTTTGAAGCTTGGCCAGAGGCTCTTGGATGTTGGTGTG 820
Db 1 AAGATTTGTGTGTGCGNAATTTGATCTTAAGGCTTGGCCA-AAGGTTCTAGATGTTGGCTGTG 59
Qy 821 GAATTTGGTGGAGGTGACTTTTACATGCGCGAGACCTTTTGAATGTTGAGGTTGTTGATTTG 880
Db 60 GCATAGTGGAGGTGACTTTTATATGCGCAAGAAATTTGAGTCTATGTTGGCATCG 119
Qy 881 ATCTCTCGTTAATATGATTTCTTTCCTTGGCCCTTGGAGGTTCTATTTGGGCTTAATTTGCTG 940
Db 120 ATCTCTCGTTAATATGATTTCTTTCCTTGGAGGCTTGGAGGCTTGGAGGCTTGGAGGCTG 179
Qy 941 TTGAGTTTGGGTAGCAGATTGCAACCAAGATAAATACCTCTGATTAATCTTTTGTATGTCA 1000
Db 180 TGGAAATTTGAAGTTGCTGATTTGCACTACAAAGACTTATCCAGATAACAGCTTTTGTATGTTA 239
Qy 1001 TCTATAGCCGTGACACCATTTCTGATATTCAGCAGAGGCTGGTGTGTTAGATCTTCT 1060
Db 240 TCTATAGCCGTGACACCATTTCTGATATTCAGCAGAGGCTGGTGTGTTAGATCTTCT 299
Qy 1061 ACAATGTTGAGGCTGAGGAGGTTAAAGTTCTAATCAGTGAATTAAGTCTGCAAGAAAGCTGTC 1120
Db 300 ACAATGTTGAGGCTGAGGAGGTTAAAGTTCTAATCAGTGAATTAAGTCTGCAAGAAAGCTGTC 359
Qy 1121 CACCTCTACCTGAAATTCGCGCTTTACATTAAGCAGAGGAGGATATCATCTCCATGATGTA 1180
Db 360 CTTCAATCGCCGAAATTTGCAAGATACATCAAGCAGAGAGGCTATGATCTCCATGATGTA 419
Qy 1181 AGGATATGGCAGATCTTAAAGATGCTGGATTTGTTGATGTTCTTCCGAGGATGAA 1240
Db 420 AAGCTTATGGCAGATCTTAAAGATGCTGGATTTGTTGATGTTCTTCCGAGGATGAA 479
Qy 1241 CTGACAGTTCAATCGAGTTCTTACGGAAGGAACTAGAGACTGTTGAGAGGAAAGGATG 1300
Db 480 CCGATCAGTTCAATCAAGTTCTGAGGCTGAAATAGCAAAAGTTGAGAGGAAAGGATG 539
Qy 1301 TGTTCATTTAGTGAATTTCTGAGGAGGATTTACATGATGATGATGATGATGATGATG 1357
Db 540 CTTCAATCAGGCTTTTCTCAGGAGAGATGATGATGATGATGATGATGATGATGATG 596

```

RESULT 12

```

US-10-021-323-8285
; Sequence 8285, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR FILING DATE: 2001-12-12
; PRIOR FILING DATE: 2001-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 8285
; LENGTH: 601

```

```

; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(601)
; OTHER INFORMATION: Clone ID: LIB3828-008-Q1-K6-B6

```

```

US-10-021-323-8285
; Query Match 22.0%; Score 353.2; DB 17; Length 601;
Best Local Similarity 75.2%; Pred. No. 1e-86;
Matches 453; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

```

QY 753 AACCAACCAAGAGTTTGTGTCATGCTGGAATCTGGAAGCTGCGCAGAGGTCCTGGATG 812
 Db 1 AACAACAACAAGATTGTGGGAGTTAGATCTTAAGCTTGGCA- AAGTCTTAGATG 59
 QY 813 TGGTGTGGATTTGGTGGAGTGACCTTTACATGGCGGAGACTTTGATGTTGAGGTTG 872
 Db 60 TGCTGTGGCAATTTGGTGGAGTGACATTTATATGGCTGAGGAATTTGATGTTGATGTT 119
 QY 873 TGGATTTGATCTCTCCGTTAATATGATTTCTTTTGGCCCTTGAGCGGTTCTATTGGGCTTAA 932
 Db 120 GGCGCAACGACCTCTCTGTAACATGATATCTTTGCTTGTGAACGAGCTACTGGACTGA 179
 QY 933 ATGCTGTGTTGAGTTTGGAGTAGCAGATTGCACCAAGATAAATCACTCCCTGATACTCTT 992
 Db 180 ATGCTGAGTGGATTTGAGTTGCTGATTCACACAGAGGTTTATCCCGACACAGTTT 239
 QY 993 TGATGTCATCTATAGCGGTGACACCATCTCTGATATTCAGGACAGCCCTGCGTTGTTAG 1052
 Db 240 TGATGTTATCTACAGCGGTGACACTATTTCTACACATTCATGACAAACCTGCACTGTTAG 299
 QY 1053 ATCTCTTACAAATGGTTGAAGCCAGGAGGTAAAGTTCTAATCAGTGAATTACTGCAAGAA 1112
 Db 300 ATCTCTTACAAATGGTTGAAGCCAGGAGGCAAACTCTCTAATGATGATTACTGCAAAAG 359
 QY 1113 AGCTGGTCCACCTCACTGATTTGGCGCTTACATTAAGCAGAGGGATATGATCTCA 1172
 Db 360 TTCCAGACTCCATCCAGGGAGTTTGTGAGTATATCAGCAGAGGCTATGATCTTCA 419
 QY 1173 TGAATGAAGGAATATGGCAGATGCTTAAAGATGCTGGATTTGTTGATGTTCTTGGCGA 1232
 Db 420 TGAATGAATATCATATGACAGATGCTTGAAGATGCTGGTTTGTGATGATGTTCTTGCAGA 479
 QY 1233 GATAGACTGAGCAGTTCTTCAAGTTCTACGAGAGGAACTAGACACTGTTGAGAGGA 1292
 Db 480 GATGCAACCGATCACTTCTACAGTTCTCGCGGTGATTCACCAAGTGGAGAGGA 539
 QY 1293 AAGAGTGTTCATGATGATTTCTGAGGAGGATTAATGACATTTGTTGGAGTTG 1352
 Db 540 GAAGGATGCAATTCATCTCTGACTTCTCTAAGGAAGACTATGATGAATAGTTGGTGGATG 599
 QY 1353 GA 1354
 Db 600 GA 601

RESULT 13

US-10-021-323-14357
 ; Sequence 14357, Application US/10021323
 ; Publication No. US20040123340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Deikman, Jill
 ; APPLICANT: Feng, Paul C.C.
 ; APPLICANT: Fincher, Karen L.
 ; APPLICANT: Ziegler, Todd E.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(52274)B
 ; CURRENT APPLICATION NUMBER: US/10/021.323
 ; CURRENT FILING DATE: 2001-12-12
 ; PRIOR APPLICATION NUMBER: US 60/255, 619
 ; PRIOR FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 17880
 ; SEQ ID NO 14357
 ; LENGTH: 525
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum

Query Match 20.6%; Score 329.4; DB 17; Length 525;
 Best Local Similarity 76.9%; Pred. No. 3.6e-80;
 Matches 402; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
 US-10-021-323-14357

QY 750 TGAACCAACCAAGAGTTTGTGTCATGCTGGAATCTGGAAGCTGCGCAGAGGTCCTGGA 809
 Db 3 TGAACAACAACAAGATTGTGGGAGTTAGATCTTAAGCTTGGCAAAAAGGTCCTAGA 62
 QY 810 TGGTGTGGATTTGGTGGAGTGACCTTTACATGGCGGAGACTTTGATGTTGAGGTT 869
 Db 63 TGGTGTGGCAATTTGGTGGAGTGACATTTATATGGCTGAGGAATTTGATGTTGATGTT 122
 QY 870 TGGTGGATTTGATCTCTCCGTTAATATGATTTCTTTTGGCCCTTGAGCGGTTCTATTGGGCT 929
 Db 123 TGTGGCAACGACCTCTCTGTAACATGATATCTTTGCTTGTGAACGAGCTACTGGACT 182
 QY 930 TAAATGTCTGTTCAGTTTGGAGTAGCAGATTGCACCAAGATAAATCACTCCCTGATAACTC 989
 Db 183 GAAATGCTCAGTGGATTTGAGTTGCTGATTCACCAAGAGGTTTATCCCGACACAG 242
 QY 990 TTTGATGTCATCTATAGCGGTGACACCATCTCTGATATTCAGGACAGCCCTGCGTTGTT 1049
 Db 243 TTTGATGTTATCTACAGCGGTGACACTATTTCTACACATTCATGACAAACCTGCACTGTT 302
 QY 1050 TAGATCTCTTACAAATGGTTGAAGCCAGGAGGTAAAGTTCTAATCAGTGAATTACTGCAA 1109
 Db 303 TAGATCTCTTACAAATGGTTGAAGCCAGGAGGCAAACTCTCTAATAAGTGAATTACTGCAA 362
 QY 1110 GAAAGCTGGTCCACCTCACTGATTTGGCGCTTACATTAAGCAGAGGGATATGATCT 1169
 Db 363 AAGTTCCAGACTCCATCCAGGGAGTTTGTGAGTATATCAGCAGAGGCTATGATCT 422
 QY 1170 CCATGATGAAGGAATATGGCAGATGCTTAAAGATGCTGGATTTGTTGATGTTCTTGGC 1229
 Db 423 TCATGATGAATATCATATGACAGATGCTTGAAGATGCTGGTTTGTGATGTTGATCTTGC 482
 QY 1230 CCAGATAGAACTGAGCAGTTCTTCCAGTTCTTACCGAGGAA 1272
 Db 483 AAGAGATCGAACCGATCACTTCTTACAGTTCTCGCGGTGAA 525

RESULT 14

US-10-021-323-17175/c
 ; Sequence 17175, Application US/10021323
 ; Publication No. US20040123340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Deikman, Jill
 ; APPLICANT: Feng, Paul C.C.
 ; APPLICANT: Fincher, Karen L.
 ; APPLICANT: Ziegler, Todd E.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(52274)B
 ; CURRENT APPLICATION NUMBER: US/10/021.323
 ; CURRENT FILING DATE: 2001-12-12
 ; PRIOR APPLICATION NUMBER: US 60/255, 619
 ; PRIOR FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 17880
 ; SEQ ID NO 17175
 ; LENGTH: 694
 ; TYPE: DNA
 ; ORGANISM: Gossypium hirsutum
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3829-031-Q6-K6-F4-PPOLYA
 US-10-021-323-17175

Query Match 20.4%; Score 326.6; DB 17; Length 694;
 Best Local Similarity 74.1%; Pred. No. 2.6e-79;
 Matches 413; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
 QY 866 AGTGTGTGATTTGATCTCTCCGTTAATATGATTTCTTTGGCTTGGAGGTTCTATTG 925
 Db 694 ATGTGTGGCAATTCATCTCTCCATATGAAATCTTTTCTCTTGAAGAGCCATG 635
 QY 926 GCTTAAATGTCTGTGAGTTGAGTAGCAGATTGCAAGATAAATACCTGATA 985

```
Db 634 CACTCGATGCTCTGTGGAATTTGCACTTGTGATGCTACACAGACTTATCCAGATA 575
Qy 986 ACTCTTTTGTGATGCTATATAGCCGCTGACACAGCTTCTGCATATTCAGACAGCTGCGT 1045
Db 574 ACAGCTTTGATGATCTATCTACAGCCGCTGACACAGATTTTACACATTTCAAGACAGCCCTGCAC 515
Qy 1046 TCTTTAGATGCTCTTACAAATGCTTGAAGCCAGAGGTAAAGTTCTTAATCAGTGAATTA 1105
Db 514 TATTTAGATATTTTACAAATGCTTGAAGCCAGGGGCAAGATTTCTTATAAGTGAATTA 455
Qy 1106 GCAAGAAAGCTGCTCCACCTCACTGAAATTCGCGCTTACATTAAGCAGAGGGGATATG 1165
Db 454 GCAAAAGCCAGAGGCTTCTGCGCCGAAATTTGCAGAGTACATCAAGCAGAGAGGCTATG 395
Qy 1166 ATCTCCATGATGAAGAAATATGGCAGATGCTTAAAGATGCTGATTTGATGATTC 1225
Db 394 ATCTCCATGATGAAGAAATATGGCAGATGCTTAAAGATGCTGATTTGATGATTC 335
Qy 1226 TTGCCAGAGATAGAACTGAGCAGTTCATTCGAGTTCTACGGAAGGAACTAGAGACTGTTG 1285
Db 334 TTGCAGAGATCGAACCCATCAGTTTATACAAAGTTCTGACAGCGTGAATTTAGACAAAGTTG 275
Qy 1286 AGAAGGAAAGATGCTGTTGATGATGATTTCTGAGGAGGATTTACATGACATTTG 1345
Db 274 AGAAGGAAAGATGCTGTTGATGATGATTTCTGAGGAGGATTTACATGACATTTG 215
Qy 1346 GAGGTTGGAATGATTAAGTTGCGGAGGACTGCGAAGGTTGAGCAACGATGGGTTCTGTTG 1405
Db 214 GGGGATGGAGCAAGCTAAATAGGACCAAGATTTGGGGAGCAGAGATGGGGTCTTTTCA 155
Qy 1406 TTGCCAAGAAAGATGA 1422
Db 154 TTGCCAACAACAACTAA 138
```

RESULT 15

```
US-10-021-323-13961
; Sequence 13961, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Peng, Paul C.C.
; APPLICANT: Pincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 13961
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(586)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3829-022-Q6-K6-H6
US-10-021-323-13961
```

```
Query Match 19.8%; Score 316.4; DB 17; Length 586;
Best Local Similarity 75.3%; Pred. No. 1:5e-76;
Matches 406; Conservative 0; Mismatches 132; Indels 1; Gaps 1;

Qy 337 TATCTTTCTGATGAAGAGTGGAATTTGGTTGAAAGATGTTGAAATGTTGAAAGCA 396
Db .. 45 TGTATCTCTCANTAAAGAGTTGCAATTTGGCGG-AGGATGCTCAAGTTGTAAGTT 103
Qy 397 GGGGTTACATTTTCTTCAGAGATCTGTTTCCATCAATCTGGGATCACAAGCCAA 456
```

```
Db 104 GGTGACATATATTTTCTTCAGGAGTCTTGCTTTCATCAATCTGGTGATTCGACAGAAAG 163
Qy 457 AGCAATCCCAACCACTACCGTGAACTTAGGTTCTACACTAAGCCCTTCAAGAGTGTAT 516
Db 164 AACAAACCAACTCACTATCGGAGCCGAGATTTTACACTAAGGTCTTTAAAGAAATGCCAA 223
Qy 517 TTGCAAGATGGATCTGGAATCTCTTATGAGCTCTCCCTACTCTAGCTGCAAAATGTAATGA 576
Db 224 GCAACTGATGATTTCTGGAATTCATTTGAATTTCTTCTGTTGGCTGCAAGTGCATGGA 283
Qy 577 GCTTATGTACAAACAAAGAAACCAAGAACAGATAGTTGTTGTCGCAAAAGTTGAT 636
Db 284 GCTTATGTAGAACAAAGAAACAGATCAGATCAGATTTGCTGGATCTTCCAAAAGTTGTT 343
Qy 637 TCTAAGGATGATAAGGGGTTCCAGCGGATTTCTGGATACTAGCCAGTACAAAGTGTAAATAGC 696
Db 344 TCAGATAATGATAAGGGATTTCCAGCGCTTCTTGGAATTTCTGTTTCACTACAAATCTAATAGC 403
Qy 697 ATTCTGCGATATGAGCGTGTATTTGGCCCTGTGTTATGTTAGCACTGGAGGATATGAACC 756
Db 404 ATTCTGCGATATGAGCGTGTATTTGGCCCTGTGTTATGTTAGCACTGGAGGATATGAACC 463
Qy 757 ACCAAAGATTTGTGTCATGCTGGACTTGAAGCCCTGGCCAGAGGTCTCTGATGTTGCT 816
Db 464 ACAAAGAAATTTGTGGGGAAGTTAGATCTTTAAGCCCTGGCCAAAGGTCTCTAGATGTTGGC 523
Qy 817 TGTGGAATTTGCTGAGGTGATCTTTTACATGCGCGGAGACCTTTGATGTTGAGGTTGTTGG 875
Db 524 TGTGGCATTTGCTGAGGTGATCTTTTATATGCTGAGGAAATTTGATGTTTCAATGTTGTTGG 582
```

Search completed: August 2, 2004, 08:38:29
Job time : 1102 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 04:40:09 ; Search time 4442 Seconds

(without alignment)
10769.746 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602

Sequence: 1 cacacccgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_esthum:*
4: em_esthum:*
5: em_esthum:*
6: em_esthum:*
7: em_esthum:*
8: em_esthum:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est5:*
15: em_estfun:*
16: em_estfun:*
17: em_estfun:*
18: em_estfun:*
19: em_estfun:*
20: em_estfun:*
21: em_estfun:*
22: em_estfun:*
23: em_estfun:*
24: em_estfun:*
25: em_estfun:*
26: em_estfun:*
27: em_estfun:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	732.8	45.7	2333	11- AY103779	AY103779 Zea mays
2	654	40.8	713	10 BE231445	BE231445 S650815 S
3	566.6	35.4	1012	11 AY105261	AY105261 Zea mays
4	538.6	33.6	1211	10 BE574969	BE574969 E07 Triph

5	465.6	29.1	1022	14	CK206655
6	464.6	29.0	767	12	BM408047
7	463.2	28.9	781	14	CF208434
8	461.6	28.8	869	14	CF513471
9	452	28.2	1115	14	CK166076
10	451.4	28.2	753	14	CB970167
11	440.6	27.5	720	14	CF508254
12	435.2	27.2	601	12	BM302134
13	429	26.8	822	10	BE040284
14	427.2	26.7	706	13	BQ120291
15	426	26.6	733	13	BQ965516
16	421.8	26.3	731	13	BQ855322
17	421.6	26.3	805	10	BE040460
18	416.4	26.0	721	14	CF482195
19	415.6	25.9	941	14	CK265554
20	413.2	25.8	682	13	BQ856500
21	410.4	25.6	703	14	CB972747
22	409	25.5	736	13	BQ852610
23	408.6	25.5	750	12	BM406709
24	407.8	25.5	718	14	CA174643
25	407	25.4	527	13	BQ587932
26	404.6	25.3	700	14	CA289955
27	403	25.2	1127	10	BE420987
28	401	25.0	770	14	CD920161
29	400.2	25.0	771	14	CA327433
30	399.2	24.9	848	14	CF513552
31	398.8	24.9	731	13	CA100170
32	398.6	24.9	806	14	CK195203
33	394.6	24.6	689	13	BQ119788
34	393	24.5	664	13	BQ871545
35	391.4	24.4	802	12	BQ573651
36	390.2	24.4	745	13	BQ762263
37	389	24.3	680	12	BI406237
38	388.6	24.3	720	12	BJ573369
39	388	24.2	672	14	CD323049
40	383.2	23.9	621	13	BU635947
41	382.4	23.9	685	13	CA133117
42	380.4	23.7	718	14	CA801290
43	380	23.7	615	13	BQ850509
44	380	23.7	655	13	CA077350
45	379.4	23.7	788	14	CA295025

ALIGNMENTS

RESULT 1	AY103779	2333 bp	linear	HTC 16-OCT-2002
LOCUS	Zea mays	PCO126548	mrna	
DEFINITION	Zea mays PCO126548 mrna sequence.			
ACCESSION	AY103779			
VERSION	AY103779.1	GI:21206857		
KEYWORDS	HTC			
SOURCE	Zea mays			
ORGANISM	Zea mays			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.			
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes			
JOURNAL	Unpublished (2002)			
REFERENCE	Coe, E.H.			
AUTHORS	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA			
TITLE	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the			

2	(bases 1 to 2333)			
3	(bases 1 to 2333)			
4	(bases 1 to 2333)			
5	(bases 1 to 2333)			
6	(bases 1 to 2333)			
7	(bases 1 to 2333)			
8	(bases 1 to 2333)			
9	(bases 1 to 2333)			
10	(bases 1 to 2333)			
11	(bases 1 to 2333)			
12	(bases 1 to 2333)			
13	(bases 1 to 2333)			
14	(bases 1 to 2333)			
15	(bases 1 to 2333)			
16	(bases 1 to 2333)			
17	(bases 1 to 2333)			
18	(bases 1 to 2333)			
19	(bases 1 to 2333)			
20	(bases 1 to 2333)			
21	(bases 1 to 2333)			
22	(bases 1 to 2333)			
23	(bases 1 to 2333)			
24	(bases 1 to 2333)			
25	(bases 1 to 2333)			
26	(bases 1 to 2333)			
27	(bases 1 to 2333)			
28	(bases 1 to 2333)			
29	(bases 1 to 2333)			

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

source

Location/Qualifiers

1. 2333

/organism="Zea mays"

/mol_type="mRNA"

/db_xref="MaizeDB:637372"

/db_xref="taxon:4577"

/clone_lib="Maize Mapping Project/DuPont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

```

Query Match      45.7%; Score 732.8; DB 11; Length 2333;
Best Local Similarity 69.8%; Pred. No. 1.5e-108;
Matches 989; Conservative 0; Mismatches 427; Indels 0; Gaps 0;

QY 1 CACACCGTGTGATTAACCATTTGAAGCTATGATCTCGATTTCTCAAGCTTCTGATCTTGAC 60
DB 488 CACTCCAAAGTGCCTCACTGTGAGTCCATGATGCTCGACTCCGCGCGCGACCTCGAC 547
QY 61 AAGAGAGAACTCTGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTCTG 120
DB 548 AAGGAGAGACGACCGAGATCTCTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 607
QY 121 GAATCTGGGCTGATTTGCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 180
DB 608 GAATCTGGGCTGATTTGCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 667
QY 181 ATTGCTCTGATTTCTCAATTTGCTGATTTCTCAATTTGCTGATTTCTCAATTTGCTGAT 240
DB 668 CTGGCACTAGACTTTTGAAGTGTGATTTGAAGTGTGATTTGAAGTGTGATTTGAAGTGT 727
QY 241 AAAATGTCAATTTATGTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTG 300
DB 728 AAGACATATACCTTCAATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTT 787
QY 301 TCATTTGATGATTTCTCAATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGAT 360
DB 788 TCTGTTGATGATTTCTCAATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTGAT 847
QY 361 AATTTGTTGAAGATTTCTCAATTTGCTGATTTGCTGATTTGCTGATTTGCTGATTTGCTG 420
DB 848 AAGCTTGTGGGAAATGTTGAATGTTGAATGTTGAATGTTGAATGTTGAATGTTGAATGTT 907
QY 421 TCTTTTTCATCATCTGGGATTCACAAAGCGAAAGGATTCACAAAGCGAAAGGATTCACAA 480
DB 908 TCATTTTTCATCATCTGGGATTCACAAAGCGAAAGGATTCACAAAGCGAAAGGATTCACAA 967
QY 481 CTAGGTTTCTACATCTAGGCTTCACAAAGGATTCACAAAGGATTCACAAAGGATTCACAA 540
DB 968 CCAAGTTTATACAGATTTTGAAGGATTCACAAAGGATTCACAAAGGATTCACAAAGGAT 1027
QY 541 TATAGCTCTCTCTACTAGTCTGAATTTGAGTCTGATTTGAGTCTGATTTGAGTCTGATTT 600
DB 1028 TTGAACTTCTCTAGTCTGAATTTGAGTCTGATTTGAGTCTGATTTGAGTCTGATTTG 1087
QY 601 CAGAACCAAGATTTGTTGTTGCAAAAGGATTTGATTTGATTTGATTTGATTTGATTTGAT 660
DB 1088 CAACACCAAGATTTGTTGTTGCAAAAGGATTTGATTTGATTTGATTTGATTTGATTTGAT 1147
QY 661 CCAATTTCTGATTTAGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 720
DB 1148 AGATTTCTGCAAGCTGCAATATCAACAGAGTGGGATTTAGTGTGATTTAGTGTGAT 1207
QY 721 GGCCCTGTTATGTTAGCACTGGAGGATTTAGAAACCAACCAAGAGTTTGTGCAATCTG 780

```

```

DB 1208 GGTGAAGGTTTGTGAGCACTGTGGAATCGAGACTACAAAGGAATTTGTGGGATGCTC 1267
QY 781 GACTTTGAAGCTCTGCGCAGAAAGTCTCGATGTTGTGTTGTTGTTGTTGTTGTTGTTG 840
DB 1268 GATCTTAAACCGGCGCAGAAAGTACTTGTGTCGGATGTGGAATTTGGAGCGCGGACTT 1327
QY 841 TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGAGTCTTCTGTTGTTGTTGTTGTTG 900
DB 1328 TACATGGCTGCAACTATGATGTTGATGTTCTTGTGTTGTTGTTGTTGTTGTTGTTG 1387
QY 901 TCTTTGCGCTTGTAGCGTTCTATTTGGCTTAAATGTGCTTGTGTTGTTGTTGTTGTTG 960
DB 1388 TCATTTGCAATTTGAACGTCGCAATTTGACGCAAGTCTCTGTTGTTGTTGTTGTTGTT 1447
QY 961 TGCAACCAAGATAAACTACTACCTGATTAATCTTTTGTGATCATCTATAGCGCTGAC 1020
DB 1448 TGCAACCAAGATAAACTACTACCTGATTAATCTTTGATCATCTATAGCGCTGAC 1507
QY 1021 CTGCATTTTACGAGCAAGCTCTGTTGTTGATGTTCTTCTACAAATGTTGTTGAGCCAG 1080
DB 1508 CTTCATACATACAGCAAGCTCTGTTGTTGATGTTCTTCTCAATTTGTTGTTGTTGTT 1567
QY 1081 GGTAAAGTCTTAATCAGTGAATTAATCTGCAAGAAAGTGTGTTGTTGTTGTTGTTGTT 1140
DB 1568 GGTAAAGTCTTAATCAGTGAATTAATCTGCAAGAAAGTGTGTTGTTGTTGTTGTTGTT 1627
QY 1141 GCTTACATTTAAGCAGAGGGATGATGATCTCATGATGTTAAGGATATATGGGAGATG 1200
DB 1628 GGTACATTTAAGCAGAGGGATGATGATCTCATGATGTTAAGGATATATGGGAGATG 1687
QY 1201 AAGATGCTGATTTGTTGATGTTCTTGTGCGAGGATAGAACTGAGCAGTTGATTCAG 1260
DB 1688 AAGATGCTGATTTGTTGATGTTCTTGTGCGAGGATAGAACTGAGCAGTTGATTCAG 1747
QY 1261 CTACGGAAGGAATAGAGACTGTTGCAAGCAAAAGCATGTGTTCAATGATGTTCTCT 1320
DB 1748 CTACGGAAGGAATAGAGACTGTTGCAAGCAAAAGCATGTGTTCAATGATGTTCTCT 1807
QY 1321 GAGGAGGATTAATGATGATGATTTGTTGAGGTTGGAATGATGTTGCGAGGATGTTG 1380
DB 1808 CAGGAGGATTAATGATGATGATTTGTTGAGGTTGGAATGATGTTGCGAGGATGTTG 1867
QY 1381 GGTGAGCAACATGGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1416
DB 1868 GGTGAGCAACATGGGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG

```

RESULT 2

BE231445

LOCUS

DEFINITION

S98015 Suaeda salsa ZAP cDNA library Suaeda maritima subsp. salsa

CDNA similar to c.elegans cDNA yk92b11.3, mRNA sequence.

BE231445

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST.

Suaeda maritima subsp. salsa

Suaeda maritima subsp. salsa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Amaranthaceae; Suaeda.

1 (bases 1 to 713)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Key Laboratory of Plant Stress Research

The Biology Department of Shandong Normal University

No. 88, Weihus East Road, Jinan, Shandong Province, 250014, PRC

Tel: (86)531-2960864

Fax: (86)531-2968954

Email: zhang@sdnu.edu.cn

Location/Qualifiers

source

1. 713

Db 1098 TACATGGCTGACGACGATGATGTTTCATGTTGTTGGCATGACCTCTCGCTCAATGATT 1157
 Qy 901 TCTTTGGCTTGAGCGTTCATGCGCTTAAATGCTGCTTCAGTTGAGCT 953
 Db 1158 TCTTTGCTCTCGAGCGTGCATGTTGCTCTCAGTGTCTGTTGAGTTGAGTT 1210

RESULT 5

CK206655 1022 bp mRNA linear EST 08-DEC-2003
 LOCUS FGAS018258 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
 DEFINITION aestivum cDNA, mRNA sequence.

ACCESSION CK206655
 VERSION CK206655.1 GI:39569045
 KEYWORDS EST.

SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.

REFERENCE

AUTHORS Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Larocque, A.,
 Links, M.G., McCarthy, B.L., Monroy, A., Muzak, I., Nilson, D.,
 Penniket, C., Roach, J.L. and Sarhan, P.
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops
 Unpublished (2003)

TITLE

JOURNAL

COMMENT

Contact: Mm L Crosby
 Bioinformatics
 University of Saskatchewan, Department of Computer Science
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,
 Saskatchewan, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas.est@usask.ca

This sequence is the direct result of the Base calling software
 Phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [9,84].

Plate: L58008 row: C column: 07.

FEATURES

source

Location/Qualifiers
 1..1022
 /organism="Triticum aestivum"
 /db_xref="taxon:4565"
 /mol_type="mRNA"
 /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
 /notes="Vector: pCMV.SPORT; Crown and developmental stages
 of spike formation in wheat cultivar Norstar. 4 mRNA
 populations were combined before constructing the library.
 The first mRNA population is from 1cm crown sections after
 30 days of cold acclimation. The second is from 1cm crown
 sections after 11 days of deacclimation (before
 deacclimation plants were fully vernalized for 49 days).
 The third is from different developmental stages of spike
 formation (5 to 50mm) that still have not emerged from the
 leaf (dissection required). The last is from different
 developmental stages of spike and seed formation after
 having emerged from the leaf (visible). First strand
 synthesis in this library was done in the presence of
 methylated dCTP thereby protecting from internal cleavage
 with NotI."

ORIGIN

Query Match 29.1%; Score 465.6; DB 14; Length 1022;
 Best Local Similarity 70.0%; Pred. No. 26-65;
 Matches 627; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

Qy 521 AAGATGGATCTGGAACTCTTATGAGCTCTCCCTACTAGTGGCAATGTTATGGAGCTT 580
 Db 25 ATGACCAAGATGGCAATCTCTTTGAGCTTCTCTGGCACTTCCCAAGTGCATTGGAGCTT 84

Qy 581 ATGTCAGAAACAAGAAACACGACGACGANTAGTTGGTTGGCCAAAAGTTGATTCTA 640
 Db 85 ATGTGAGAGCAGAGAACGACGACGATATGTTGGCTATATGGGAGAGGTCAGATGCA 144
 Qy 641 AGATGATPAGGGGTTCCAGCGATTTCTGGATACTAGCCAGTACAAGTGTAAATAGCATT 700
 Db 145 CAGAGACAAAGGCTTTTCAGAGATTCTGACAAATGTGCAATACAAATCCATCGAATCC 204
 Qy 701 TGGCATATGAGCGTGTATTTGGCCCTGTTATGTATAGCACTGGAGGATATGAACCA 760
 Db 205 TGGGTTATGACCGCTGTTTGGGNGGGTTATGTCAGACCGGTGGATTCGAGACACAA 264
 Qy 761 AAGAGTTTGTGTCATGCTGAGCTTGAAGCTGGCCAGAGGTCTCTGAGTTGTTGTG 820
 Db 265 AAGAAATTTGTGACAGCTGTGATCTGAAGCTGGCCAGAGGTGCTCGAGTTGTGGTGG 324
 Qy 821 GAATTTGGTGGAGGTGACTTTTACATGCGGAGACCTTTTATGTTGAGGTTGTTGATTTG 880
 Db 325 GCATCGGAGGAGGTGACTTTCTACATGCTGAAACCTATGATGTCATGTCCTCGGCATCG 384
 Qy 881 ATCTCTCCGTTAAATGATTTTCCTTTGGCCCTTGAGGCTTCTATTTGGGCTTAAATGCTG 940
 Db 395 ATCTTCCATCAACATGTTTCTGTCGATCGAGGTGCCATCGGGCTCTGCTCTCGG 444
 Qy 941 TTGAGTTTGGAGTGAAGATGTCACCAAGATAAATACCTCTGATTAATCTTTTGTGATGTC 1000
 Db 445 TTGAGTTTGGAGTTGCTGACTGCAACCAACCAAGGAATACGACAGAACACGTTTGTATGA 504
 Qy 1001 TCTATAGCCGTGACACCAATCTGCAVATTTCAGGACAAAGCTCGCTTGTGTTAGTCTTCT 1060
 Db 505 TCTACAGCCGTGACACCAATCTCTCCATTCAGACAAACCTGCTCTGTTGAGAACTTCT 564
 Qy 1061 ACAATAGTTTGAAGCAGAGGTAAAGTTCTAATCAGTATGTTACTGCAAGAAAGCTGCTC 1120
 Db 565 TCAGTGGCTCAAGCTTGGGGGCAAGTGTGATCAGTGACTACTGCGAGGAGCCCTGGGA 624
 Qy 1121 CACCTCACTGAATTCGCCCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAA 1180
 Db 625 CACCTCAGAGGAATTTGCTGATACATCAGCAGAGAGGCTATGACCTCCATGACGTGA 684
 Qy 1181 AGGAATATGGCAGATGCTTAAAGATGCTGGATTTGTTGATGTTCTTCCGAGGTAGAA 1240
 Db 685 AGACCTATGGAAGATGCTGAGATGCCGTTTCCAGTGTGCTGCCGAGAGACCGCA 744
 Qy 1241 CTGAGCAGTTCATTCGAGTTCTACGGAAGAACTAGACACTCTTTCAGAGGAAAGAGATG 1300
 Db 745 CGACCAAGTTCTCGACGCTCTCGAGAGGGAGCTGGCGAGAACGAGAAACAGGAGG 804
 Qy 1301 TGTTCAATAGTATTTCTCTGAGGAGGATTTACAATGACATTTGAGGTTGGATGATA 1360
 Db 805 CTTCTCGGGGACTTCACCCAGGAGGACTACGACGACATGTCACCGCTGGAGCGGA 864
 Qy 1361 AGTTGGCGGAGCTGCCAAGGTGAGCAACGATGGGCTGTTGTTGTTCCGCAAGAG 1416
 Db 865 AGCTGAAGCGAGCTCCCGCGGCGAGCAGAGTGGGGGCTGTTTCATCGCAACAGG 920

RESULT 6

BM408047

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BM408047 767 bp mRNA linear EST 10-MAR-2003
 EST582374 potato roots Solanum tuberosum cDNA clone CPRO33G1 5'
 end, mRNA sequence.
 BM408047
 BM408047.1 GI:18259677
 EST.
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
 Asterids; Lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 767)
 van der Hoeven, R., Sun, H., Karameycheva, S.A., Tsai, J., Van Aken, S.,
 Utterback, T., Chiemingo, A., Bougri, O., Buell, C.R., Renning, C.,

Tankley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 CONTACT: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3.

FEATURES

Location/Qualifiers
 1..767
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="CPRO33G1"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Cornell University, Tanksley lab;
 sequencing: The Institute for Genomic Research. Roots were
 isolated from in vitro grown stem cuttings on CM medium.
 roots were isolated two weeks after placing the stem
 cuttings from in vitro grown plants on medium."

ORIGIN

Query Match 29.0%; Score 464.6; DB 12; Length 767;
 Best Local Similarity 75.4%; Pred. No. 3.3e-65;
 Matches 578; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY 439 GGGGATCACAAGCGAAGCAATCCACACCTACCGTCACTAGGTTCTACACTAG 498
 DB 1 GGGACCAACAGCAAGCAACCCACCATTTTCGGAGCTAGATTACACAAAG 50
 QY 499 GCCTTCAGAGATGTCATTGCAAGATGATCTCGAAATCTTATGAGCTCTCCCTACT 558
 DB 61 GTCTTTAAAGATGTCATATAAATGCTGTGATGTAATCATTTGAACITTTCTCTCATC 120
 QY 559 AGCTGCAATGTTAGGAGCTTATGTCAGAAACAGAAACAGAAACAGAAATGATGAG 618
 DB 121 GGTTCAGATGCTTGGCTTATGTTAAACAAAGAAATCAGAAATGATGTTG 180
 QY 619 TTGTGCAAAAGTTCATAGGATGATTAAGGCTTCCAGGATTCAGGATTCGGATCTAGC 678
 DB 181 CTATGCAAAAGGTTAAATCTGAGGATGACAGGAGATTCAGGCTTTCTCGCAATGTT 240
 QY 679 CAGTACAGTGTAAATAGCAATCTCGCATATGAGCGTGTATTTGGCCCTGTTATGTAGC 738
 DB 241 CAATACAAATGTAGTGGCATCTGCGATATGAACGTGCTTTGGAGAAAGTTATGTGAGC 300
 QY 739 ACTGGAGGATATGAACCAACCAAGATTTGTGTCAATGCTGAGCTTGAAGCTCGCCAG 798
 DB 301 ACAGAGGACTGTGATACCAACAAAGATTCGTTCTATGTGATCTTCAACCTGCGCA 360
 QY 799 AAGTCTGAGATGTTGGTGTGGAATTTGTGTGAGGTGATCTTTACATGCGGAGACTTT 858
 DB 361 AAGTCTGAGATGTTGGTGTGGAATTTGTGTGAGGTGATCTTTACATGCGTGAATAT 420
 QY 859 GATGTTGAGATGTTGGTGTGGAATTTGTGTGAGGTGATCTTTACATGCGTGAATAT 918
 DB 421 GATGTTGAGATGTTGGTGTGGAATTTGTGTGAGGTGATCTTTACATGCGTGAATAT 480
 QY 919 TCTATGGGCTTAAATGCTGCTGATGTTGAGGTGAGGATTCACCAAGATAAATCTAC 978
 DB 481 TCAATGGTCTCAATGTCGCTGATGTTGAGGTGAGGATTCACCAAGATAAATCTAC 540
 QY 979 CCTGATAATCTTTTGAATGTCATCTATAGCGTGAACCATTTCTGCAATATTCAGGACAAG 1038
 DB 541 CCTGATGGCAATTTGATGTGATTTACAGTCGAGACATATCTCTTCATCTCCAGACAA 600

QY 1039 CTGCGTGTGTTTACATCTCTTACAAATGTTTGAAGCCAGAGGTTAAAGTTCTTAATCAGT 1098
 DB 601 CCCAGTTTATTCAGATCTTTTACAGTGGCTGAAACAGGAGGCAAGTCTCTATAAGT 660
 QY 1099 GATTACTGCAAGAAGCTGGTCCACCTGACCTGATTCGCGCTTACATTAAAGCAGAGG 1158
 DB 661 GATTACTGCAATATGCCAATACACAGCATCGAGAGTGTTCGATATATAAGCAAGG 720
 QY 1159 GGATATGATCTCCATGATGTAAAGGAATATGGCAGATGCTTAAAGA 1205
 DB 721 GGTATGATTTACATGATGTTGCAACATATGGCCAGATGCTCAAGA 767

RESULT 7

CF208434 781 bp mRNA linear EST 01-AUG-2003
 CAB20003_IIa_Fa_G05 Cabernet Sauvignon Flower bloom - CAB2 Vitis
 vinifera cDNA clone CAB20003_IIa_Fa_G05 5', mRNA sequence.
 CF208434
 CF208434.1 GI:33402807
 EST.
 Vitis vinifera
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 REFERENCE 1 (bases 1 to 781)
 Goes da Silva, P., Iandolo, A., Lim, H., Baek, J., Jones, K. and
 Cook, D.
 Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 CONTACT: Douglas Cook, PhD
 UC Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seq primer: ACGGTACCGGACATATGCC.

FEATURES

Location/Qualifiers
 1..781
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB20003_IIa_Fa_G05"
 /sex="Hermaphrodite"
 /dev_stage="Bloom"
 /clone_lib="Cabernet Sauvignon Flower bloom - CAB2"
 /note="Organ: Flower - Bloom; Vector: pDNR; Site 1: SfiI;
 Site 2: SfiI; CAB2 is a cDNA library of Vitis vinifera cv.
 'Cabernet Sauvignon' Clone 8 berries. Samples were
 collected at full bloom (80 to 100% flowers showing
 dehiscence of calyxtraps or caps and anthers fully
 extended). Sampled vines were located at the University of
 California, Davis, Experimental Vineyard. cDNAs were made
 by oligo-dT priming and directionally cloned. 5' and 3'
 adaptors were used in cloning as follows:
 5'-AACGATGTTGATCAACGAGAGTGGCTTACGCGCCGG-3' and
 5'-ATTCTAGAGCGGAGCGCGGACATG-dt(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 28.9%; Score 463.2; DB 14; Length 781;
 Best Local Similarity 74.6%; Pred. No. 5.4e-65;
 Matches 582; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 107 GAATAATGCTCTTGGAACTTGGCGTGTATGCTGTTTACTGCTGATTCGCTGATG 166
 DB 2 GGAAATATGCTGCTGGAATAGGGGAGGAATTTGGCGTTTCACAGCGGATGATGCTAGC 61

Db 613 TCTGAATCTCCTCCTCATCCTCAATGTTGGAGCTTATGTGAGAAACAAAGAAT 672
 Qy 601 CAGAACCAATAGTCTGTTGCGCAAAAGTTGATCTTAAGCATGTAAGGGTTCCAG 660
 Db 673 CAAACCAAGATCTATGTTATGTCAGATGCTTAATTCACAAATGACAAAGGGTTCCA 732
 Qy 661 CGATTTCTGGATCTACGCCAGTACAGTGAATAGCATCTCGGATATGAGGCTGTATTT 720
 Db 733 CGTTTCTCGGATATGACAGTAAATGACGGGGCATATTACGCTATGAAAGTGTCTTT 792
 Qy 721 GGCCTCGTGTATGATGACCTGAGGATATGAACCAACCAAGAGTTGTGTCAATGCTG 780
 Db 793 GGTGAAGGCTTGTGTGATGACAGGGGACTTGAACCAACCAAGAAATTTGTGACAAAATTG 852
 Qy 781 GACTTGAAGCTGGCCCA 797
 Db 853 GACCTTAAGCCTAGCCA 869

RESULT 9
 CK166076/c
 LOCUS
 DEFINITION 1115 bp mRNA linear EST 05-DEC-2003
 FGAS050129 Triticum aestivum FGAS: Talt7 Triticum aestivum cDNA,
 mRNA sequence.

ACCESSION CK166076
 VERSION CK166076.1 GI:38998769

KEYWORDS EST.
 SOURCE Triticum aestivum (bread wheat)

ORGANISM

REFERENCE 1 (bases 1 to 1115)
 AUTHORS Allard, F., Crosby, W.L., Danyluk, J., Budes, P., Frick, M., Gaudet, D.,
 Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
 Linke, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
 Penikiet, C., Roach, J.L. and Sarhan, F.
 Functional Genomics of Abiotic Stress in Wheat and Canola Crops
 Unpublished (2003)
 Contact: Wm L Crosby
 Bioinformatics

TITLE University of Saskatchewan, Department of Computer Science
 JOURNAL 1C101 Engineering Building, 57 Campus Drive, Saskatoon,
 COMMENT Saskatoon, S7N 5A9, Canada
 Tel: 306 966 1769
 Fax: 306 966 2033
 Email: fgas_estecs.usask.ca

Phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [65,863].
 Plate: Talt707 row: K column: 02.

FEATURES

Source 1..1115
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Wheat line CI 14106"
 /db_xref="taxon:4565"
 /lab_host="DH5 alpha"
 /clone_lib="Triticum aestivum FGAS: Talt7"
 /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
 subtractive hybridization) cDNA library from genotype
 CI14106 cold hardened at 2 C for 21 days and 49 days
 (equal amount of cDNA pooled together before subtraction,
 teater) and subtracted against genotype CI14106
 non-hardened (20 C) (driver). Nitro-pyrole anchored
 oligo-dt priming and non-directional cloning."

ORIGIN

Query Match 28.2%; Score 452; DB 14; Length 1115;
 Best Local Similarity 68.9%; Pred. No. 2.9e-63;

RESULT 10
 CB970167

Matches 663; Conservative 0; Mismatches 295; Indels 4; Gaps 3;
 Qy 455 AAGCAATCCCAACCCACTACCGTGAACCTAGGTTCTACACTAAGGCTTCAAGAGTGC 514
 Db 1103 AAGTAACCCGACCCACTACCGGAGCGAGGTTTACCAACCAAGGTGTAAAGGA--TGC 1046
 Qy 515 ATTTCGAGATGATCTCGAAACCTCTTATGATGCTCTCCCTACTAGCTGCAATGATTG 574
 Db 1045 CACTCTTAGACCAAGAGGGATTCCTTGAAGCT--TCCTTCGTAACCTCCAGTGCATG 987
 Qy 575 GAGCTTATGTGAGAAACAAAGAAACCAAGAACCAAGATAGTGTGTGGCAAAAGTTG 634
 Db 986 GAGCTTATGTGAAAGCAAGGAGAACCCAGACCCAGATATGTGTGGCTGTGGGAGAGGTCA 927
 Qy 635 ATTCTAAGGATGATAAGCGGTTCCAGCGATTTCGATACTAGCCAGTACCAAGTGTATAA 694
 Db 926 AGTCCACAGAGACAAAGGCTTTCAGAGATCCCTGCAATGTGCAGTACAAATCCACTG 867
 Qy 695 GCATCTCGATATGAGCGGTATTTGGCCCTGTTATGTTAGCACTGGAGGATATGAA 754
 Db 866 GAATCTTGGTTATGAGCGGT--GTTGGGAGGGTTATGTTAGCACCGGTGATTCGAGA 808
 Qy 755 CCACAAAGAGTTTGTGTCAATCTGACTGAAAGCTGAGCCAGAGGTCCTGGAGTTG 814
 Db 807 CCACAAAGAGTTTGTGCAAAAGCTGCAAGCTGTAAGCTGTCAGAAAGTGTCTCGATAG 748
 Qy 815 GTTGTGAATTTGGTGGAGGTGACTTTTACATGCGGAGGACCTTTGATGTTGAGGTTGTTG 874
 Db 747 GGTGTGTTATCGAGGAGGGGACTTCTACATGGCTGAACCTATGATGATTCCTGCTCG 688
 Qy 875 GATTTGATCTCTCGGTTAATGATGATTTCTTGGCCCTTGAAGCTTCTATTTGGGCTTAAT 934
 Db 687 GCATCGATCTTCTATCAACATGGTTTGGTTCCAAATGAGCGTGCATCGGGCGCTCGT 628
 Qy 935 GTGCTGTGAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 994
 Db 627 GCTCGGTTGAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 568
 Qy 995 ATGTCATCTATAGCGGTGACACCACTTCTGCATATTCAGGACAGGCTGCTGTTTGTAGAT 1054
 Db 567 ATGTCATCTATAGCGGTGACACCACTTCTGCATATTCAGGACAGGCTGCTGTTTGTAGAT 508
 Qy 1055 CCTTCTACAAATGTTGAAGCCAGAGGTAAAGTTCTTAATCAGTGAATTAATCTGCAAGAA 1114
 Db 507 ACTTCTTCAAGTGGCTCAAGGCTGGGGGCAAGTGTGATCACTGACTTACTGCAAGAGCC 448
 Qy 1115 CTGGTCCACCTCCTCCTCAATTTCCGCGCTTACATTAAGCAGAGGGGATATGATCTCCATG 1174
 Db 447 CTGGGACACCATCAGAGCAATTCGCTGTGTACATCAAGCAGAGAGGCTATGACCTCCATG 388
 Qy 1175 ATGTAAGGATATGAGGAGATGCTTAAAGATGCTGAATTTGATGTTGTTGCTGCGAGG 1234
 Db 387 ATGTGAAGACCTTACGAAAGATGCTTGAAGATGCGGTTTCCATGATGCTGCTGAGGG 328
 Qy 1235 ATAGAATCTGAGCAGTTCAATCGAGTTCTACGAAAGGAATCTAGAGACTGTTGAGAGGAAA 1294
 Db 327 ACCGCAACCGACCGTCTCTGAGGCTCTCGAGAGGGAGCTGGGCGAGAGCCGAGAGAAC 268
 Qy 1295 AGATGTTGTTCAATAGTGAATTTCTGAGGAGGATTAACATGACATTTGAGAGTTGCA 1354
 Db 267 AGAGGCTTCTCTGCGGACTTCAACCCAGGAGGACTACGAGACATGCTACACGCGCGGA 208
 Qy 1355 ATGATAAGTTGCGGAGGACTCCCAAGGAGGTGAGCAACGATGGGTCTGTTGCTGTTGCCA 1414
 Db 207 GCGCGAAGTTGAAGCGGAGCTCTGCCCGGAGCAGAAAGTGGGGGCTGTTTTCATCGCAGCA 148
 Qy 1415 AG 1416
 Db 147 AG 146

LOCUS CB970167 753 bp mRNA linear EST 30-APR-2003
 DEFINITION CAB10003_11a_Fa_H06 Cabernet Sauvignon Flower Pre-bloom - CAB1
 Vitis vinifera cDNA clone CAB10003_11a_Fa_H06 5', mRNA sequence.
 ACCESSION CB970167
 VERSION CB970167.1 GI:30253616
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 753)
 Goes da Silva, P., Iandolino, A., Lim, H., Baek, J., Jones, K. and
 Cook, D.
 Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 CONTACT: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcook@ucdavis.edu
 Seq primer: ACGGTACCGACATATGCC.
 Location/Qualifiers
 1..753
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAB10003_11a_Fa_H06"
 /sex="Hermaphrodite"
 /dev_stage="Pre-bloom"
 /lab_host="DH5alpha"
 /clone_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"
 /note="Organ: Flower - Pre-bloom; Vector: pDMR; Site: 1;
 SfiI; Site 2: SfiI; CAB1 is a cDNA library of Vitis
 vinifera cv. 'Cabernet Sauvignon' Clone 8 flowers. Samples
 were collected approximately eleven days before onset of
 bloom (clusters at this stage were fully developed and
 flowers with calypters or caps still attached. Sampled
 vines were located at the University of California, Davis,
 Experimental Vineyard. cDNAs were made by oligo-dT priming
 and directionally cloned. 5' and 3' adaptors were used in
 cloning as follows:
 5'-AAGCAGTGATCAACGACAGTGGCCATACGGCGGG-3' and
 5'-ATTCTAGCGCGGCGGCGGACATG-3' (30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

Query Match 28.2%; Score 451.4; DB 14; Length 753;
 Best Local Similarity 77.2%; Pred. No. 4.4e-63;
 Matches 548; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 714 TGTATTTCGCGCTGTTAGTCTAGCCTGAGGATATGAACCAACCAAGAGTTTGTC 773
 Db 1 TGTCTTTCGAGAGGCTTTGAGCAGGAGGATTTGAGACAAAGAAATTTGGG 60
 QY 774 AATCTGCACTTGAAGCTGCGCAGAGGCTCTGGATGTTGGTTGGAATTTGGTGGG 833
 Db 61 AAAACTGGACCTTAAACCTGCGCAGAGGCTCTAGATGCGCTGTGGCATTTGGAGG 120
 QY 834 TGACTTTTATACGCGGAGACCTTTGATGTTGAGTGTGTTGATTTGATCTCTCGTTAA 893
 Db 121 TGACTTCTATATGCGAGGACCTTTGATGTTGAGTGTGTTGATTTGATCTCTCGTTAA 180
 QY 894 TATGATTTCTTTCGCTGAGGCTTTATTTGGCTTAAATGCTGTGTTGAGTTGAGGT 953
 Db 181 TATGTTTCTTTCGCTGAGGCTTTATTTGGCTTAAATGCTGTGTTGAGTTGAGGT 240
 QY 954 AGCAGATTGCACCAAGATAAACTACCTGCTGATTAATCTTTTGTATGTCATCTATAGCCGTGA 1013

Db 241 AGCTGACTGCACCAAGAAATCATATCTCTGACACACATTTGATGTATCTACAGCGTGA 300
 QY 1014 CACCAATTCGATATTTAGGACCAAGCTGCTGTTTGTAGATCTCTACAAATGGTTGAA 1073
 Db 301 CACCAATTCGATATTTAGGACCAAGCTGCTGTTTGTAGATCTCTACAAATGGTTGAA 360
 QY 1074 GCCAGAGGTAAGTTCTTAATCACTGATTAATCTGCAAGAAAGCTGTCACCTCCACTGA 1133
 Db 361 GCCAGAGGTAAGTTCTGATTTAGTATTAATCTGCAAGAAAGCTGTCACCTCCACTGA 420
 QY 1134 ATTGCCCCCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAAGAAATATGGGA 1193
 Db 421 ATTTCAGAAATATATTAACCAAGAGGGATGATGATCTTCACGATGTGGAAGCATATGGTGA 480
 QY 1194 GATGCTTAAAGATGCTGGATTTGTTGATGTTCTTCCGAGGATAGAACTGACGAGTTCAAT 1253
 Db 481 GATGCTTAGAGATGCTGTTTCAATTCAGGTCATTTGAGAGGATAGAACTGATCACTTCT 540
 QY 1254 TCGAGTCTTACGGAAGGAACCTAGAGACTGTTGAGAGGAAAGAGATGTTTCAATTTAGTGA 1313
 Db 541 ACAAGTCTTCGAGAGGGGATTAATGCTGTTGAGACAAACAAGAAAGATTTGTTTCAGGA 600
 QY 1314 TTTCTCTGAGAGGATTAACAATGACATTTGTTGAGGTTGGAATGATAGTTGCGGAGGAC 1373
 Db 601 CTTCTCTGAAGACCACTACAAATGAAATAGTTGATGTTGGAAGGCTAGCTGATGAGGAG 660
 QY 1374 TGCCAAAGGCTGACCAAGCATGGGCTCTGTTGTTGTTGCCAAGAAAGATGAA 1423
 Db 661 TTCACTGCTGAGCAGAGATGGGGCTTTTCATTTGCAAGAAGAGTGA 710

RESULT 11
 LOCUS CB970167
 DEFINITION CB970167 753 bp mRNA linear EST 08-SEP-2003
 Sweet Orange (Citrus sinensis (L.) Osbeck) Citrus sinensis cDNA
 clone MVP-65_G10 5', mRNA sequence.
 ACCESSION CB970167
 VERSION CB970167.1 GI:34523438
 KEYWORDS EST.
 SOURCE Citrus sinensis
 ORGANISM Citrus sinensis
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosid II; Sapindales; Rutaceae; Citrus.
 1 (bases 1 to 720)
 Chaparro, J.X., Bausher, M.G., Dang, P., Hunter, W.B., McKenzie, C.L.,
 Niedz, R.P., and Shatters, R.G. Jr.
 Expressed Sequence Tags from Immature Ovaries of Valencia Sweet
 Orange (Citrus sinensis (L.) Osbeck)
 Unpublished (2003)
 CONTACT: Chaparro, JX
 Horticulture and Breeding Research Unit
 USDA, ARS, U. S. Horticultural Research Laboratory
 2001, South Rock Road, Fort Pierce, FL 34945, USA
 Tel: 772 462 5830
 Fax: 772 462 5986
 Email: jchaparro@usda.gov
 Seq primer: T3 Primer.
 Location/Qualifiers
 1..720
 /organism="Citrus sinensis"
 /mol_type="mRNA"
 /cultivar="Valencia"
 /db_xref="taxon:2711"
 /clone="MVP-65_G10"
 /tissue_type="Immature ovaries"
 /clone_lib="Immature Ovaries from field-collected Valencia
 Sweet Orange (Citrus sinensis (L.) Osbeck)"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; Standard library construction protocols from
 Stratagene cDNA synthesis kit (cat.# 200401-5) and Uni-ZAP


```

Db      541 TGTGACACAGAGGAGTTCACAAACAAAGAGTTGTGTCGAGCTTGACCTGAGCC 600
QY      792 T 792
Db      601 T 601

RESULT 13
LOCUS   BE040284
DEFINITION OEOJA05 OE Oryza sativa cDNA 5' similar to
          s-adenosyl-methionine-sterol-c-methyltransferase, mRNA sequence.
ACCESSION BE040284
VERSION   BE040284.1 GI:8335300
KEYWORDS EST.
SOURCE   Oryza sativa
ORGANISM Oryza sativa
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Ehrhartoideae; Oryzaceae; Oryza.
          1 (bases 1 to 822)
          Bohnert, H.J., Borchert, C., Brazil, S., Brooks, J., Eaton, M.,
          Ferreira, H., Kawasaki, S., McCollough, A., Michalowski, C.B.,
          Palacio, C., Scara, G., Wheeler, M. and Zepeda, G.R.
          Functional Genomics of Plant Stress Tolerance
          Unpublished (2000)
          Contact: Michalowski, C.B.
          University of Arizona
          Bio Sciences West room 513, Tucson, AZ 85721, USA
          Tel: 520-621-7982
          Fax: 520-621-1697
          Email: cbm@u.arizona.edu
          An open reading frame exists.
          Insert Length: 1 Std Error: 0.00.
          Location/Qualifiers
            1..822
              /organism="Oryza sativa"
              /mol_type="mRNA"
              /strain="Pohkali"
              /db_xref="taxon:4530"
              /tissue_type="roots"
              /dev_stage="1 week"
              /clone_lib="OG"
              /note="2-3 d 150mM NaCl"

ORIGIN
Query Match      26.88; Score 429; DB 10; Length 822;
Best Local Similarity 71.74; Pred. NO. 1.7e-59;
Matches 589; Conservative 0; Mismatches 230; Indels 2; Gaps 2;

QY      475 CGTCAACCTAGTTCTACACTAAGGCTTCAAGAGTGTCAATTTGCAAGATGGATCTGGA 534
Db      1 CGGAGGCAAGGTTTGAGACTAAGGTGTTAAAGAGTGTCAAGCTTTGATCAAGATGG 60
QY      535 AACTCTTATGAGCTCTCCCTACTTACTGCAATGTATTGGAGCTTATGTCAGAAACAG 594
Db      61 AATCTCTTGAACTCTCTGACTTACTTCTCAAGTGTGTGAGCTTACGTGAAAGCAAG 120
QY      595 AAAACACGACCAAGTATTGTTGGTGTGTCGCAAAAGTTGATTTCTAAGGATGATAGGGG 654
Db      121 AAAAATCAAAACCAAGATATGTGGCTATGSCAAAGAGTTGATTTCAACAGAGATCGGGG 180
QY      655 TTCAGGAGTTCTGGAATCTAGCCAGTACAGTGTATAGATCTCGGTATATGCGGT 714
Db      181 TTTCAAGATTTTGGACATGTGCAATCAAGCAGTGGATATACGCTATGAAACG 240
QY      715 GTATTGGCCCTGTTATGTTAGCTAGGAGGATGAAACCAACAAAGATTTGTGTCA 774
Db      241 ATCTTTGAGAGAGCTTGTGTGACACTGTGTGCAATTTGAACTACAAAGATTTGTGAC 300
QY      775 ATGCTGCACTTGAAGCTCGCCAGAGGTCCTCGATGTTGGTGTGGATTTGGTGAGGT 834

```

```

Db      301 AGGCTGGATCTCAAACTGGCCGAGAGCTTCTTGATGTTGGATGTGGAATTTGGGGCGGT 360
QY      835 GACTTTACATGGCGAGACCTTTGATGTTGAGTTGTTGAGTTTGAATCTCTCCCTTAAT 894
Db      361 GATTTTATATGGCTGACAAAGTATGATGTTCAATGTTGTTGTTGATGATCTTTGATAAAC 420
QY      895 ATGATTTCTTTGCTTGGCCCTTGAGCGCTTCTATTGGGCTTTAAATGTGCTGTTGAGTTTGAAGTA 954
Db      421 ATGGTTTCTTTTGCACCTTGAGCGGTGCTATTGGGCGTAAAGTCTCAGTTGAGTTTGAAGTC 480
QY      955 GCAGATTGACCAAGATAAACTACCTCGATAAATCTTTTGTGATGTCATCTATAGCCGTCAC 1014
Db      481 GCTGATTGCAACCAAAAGACATACCCAGACACACGTTTGGCTGATCTATACAGTCTGTGAT 540
QY      1015 ACCATTCTGCAATATCAGGACAGCGCTGCTGTTGTTAGATCTTCTACAAATGTTGAAG 1074
Db      541 ACTATCTCTACATACAGATAAAACCTCCTCACTATTAAAGTTTCTTCAAGTGGCTCAA 600
QY      1075 CCAGAGGTAAAGTTCTTAATCAGTGATTAATGCAAGAAAGCTGTGTCACCTCACCTGAA 1134
Db      601 CTGGGGGTAAAGTCTTAATTTAGTATTAATGCAAGTGTGCTGGGAAACCTTCAGAGAG 660
QY      1135 TTCCGCCCTTACATTAAGCAGAGCGGATGATGATCTCCATGATGTAAAGGAATATGGGAG 1194
Db      661 TTCCGAGCTTACATTAAGCAAAAGGGTTATGACCTTAAACCAAGTCAAGGGCTTACGGAG 720
QY      1195 ATGCTTAAAGATGCTGGATTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCAAT 1254
Db      721 ATGCTTCAA-ATGCTGGTTTCCATTGATGATGTTGTTGAGACCGCACCGATCATCTC 778
QY      1255 CGAGTTCTAGGAGGAACTAGAGACTGTTGGAAGAGGAAA 1295
Db      779 CGAGTTCTAGAGAGGAGCTGTGTAAGATTGAAGACCAAAA 819

RESULT 14
LOCUS   BQ120291
DEFINITION EST605867 mixed potato tissues Solanum tuberosum cDNA clone STMEP82
          5' end, mRNA sequence.
ACCESSION BQ120291
VERSION   BQ120291.2 GI:21919928
KEYWORDS EST.
SOURCE   Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
          asterids; lamids; Solanales; Solanaceae; Solanum.
          1 (bases 1 to 706)
          Buell, C.R., Hart, A., Baker, B., Tankale, S., Fry, W., Smart, C.,
          Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
          Karamycheva, S.A.
          Generation of a set of potato cDNA clones for microarray analyses
          Unpublished (2002)
          On Apr 17, 2002 this sequence version replaced gi:20172253.
          Contact: Robin Buell
          The Institute for Genomic Research
          9712 Medical Center Dr, Rockville, MD 20850, USA
          Email: potato-array@igrr.org
          This clone can be obtained from the University of Arizona Genomics
          Institute. Orders can be made through URL:
          http://genome.arizona.edu/orders/
          Seq primer: 13.
          Location/Qualifiers
            1..706
              /organism="Solanum tuberosum"
              /mol_type="mRNA"
              /cultivat="Kenuebac or Binjete"
              /db_xref="taxon:4113"
              /clone="STMEP82"
              /tissue_type="mixed tissues"
              /lab_host="SOLR"
              /clone_lib="mixed potato tissues"

```


ORIGIN

Query Match 26.7%; Score 427.2; DB 13; Length 706;
Best Local Similarity 75.4%; Pred. No. 3.5e-59;
Matches 531; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 464 CCACCACTACCGTGAACCTAGGTTCTACATCAAGCCCTTCAAGAGTGTCATTGCAAG 523
DB 2 CAACCCATTATCGGAGCCCTAGATTCTTACACAAAGGTGTTAAAGAAATGTCATATAATG 61

QY 524 ATGATCTGGAATCTTATAGCTCTCCCTACTAGCTGCAAAATGATTCGACTTATG 583
DB 62 CTGTGTGATGTAATCATTTTGAATCTTCTCATGTGTCAGTGTCATTCGAGCTTATG 121

QY 584 TCAGAAACAAAGAAACCAAGACCATAGTTAGTTGTTGTCGCAAAAGTTGATCTAAG 643
DB 122 TTAAGAAACAAAGAAATCAGATCAGATTTGTTGGCTATGTCGCAAAAGTTAATCTGAG 181

QY 644 ATGATAGAGGTTCCACGATTTCTGCACTAGCAGTACAGTCAAGTGTATAGCTTCTGC 703
DB 182 ATGACAGGGATTCACGGTTCTTGCAAAATGTTCAATACAAATGATGTGCACTATGC 241

QY 704 GATATGACGTTATTTGGCCCTGTTATGATGACCTGAGGATGATCAACCAAG 763
DB 242 GATATGACGTTCTTTGGAGAGGTTATGTGACGACGAGGACTTATACCAAAAG 301

QY 764 AGTTTGTCTCAATCTGCACTTGAAGCTGCGCAGAGGTCTCTGGATGTGTTGTGAA 823
DB 302 AATGCTTCTATGTTGCACTCTCAACCTGCGCAAAAGTCTTGTGCTGTGAA 361

QY 824 TTGTGAGGAGTCTTTATGACGCGGAGACCTTATGATGAGGTTGTTGATTTGATC 883
DB 362 TAGTGTGAGGAGTCTTTATGACGCGGAGTATGATGATGATGATGATGATGATGATG 421

QY 884 TCTCGTTTATATGATTTCTTGGCCCTTGCAGGCTTCTATGCGCTTAAATGTGCTGTG 943
DB 422 TCTCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481

QY 944 AGTTTGTGAGTACGATGTCACCAAGATAAATCACTGATGATGATGATGATGATGATG 1003
DB 482 AATTTGAGGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541

QY 1004 ATAGCGGTGACACCATCTGCAATTCAGGACGAGCTGCTGTTGATGATGATGATGATG 1063
DB 542 ACAGTGTGAGACATATCTTCAATCCAGACCAACCCAGCTTATTCAGATCTTCTTACA 601

QY 1064 AATGTTTGAAGCCAGGAGTAAAGTTCTAATCAAGTATGATGATGATGATGATGATGAT 1123
DB 602 GGTGCTGAAACAGGAGGCAAGTCTCTAAGTATGATGATGATGATGATGATGATGATG 661

QY 1124 CCTCACCTCAATTCGCGCTTACATTAAGCAGAGGGGATGAT 1167
DB 662 CATCGACAAAGTTTCTGATGATTAAGCAAGGGGTTATGAT 705

RESULT 15
BQ965516/c
LOCUS
DEFINITION
BQ965516
VERSION
BQ965516.1
KEYWORDS
SOURCE
ORGANISM

733 bp mRNA linear EST 21-AUG-2002
clone QHB22C02, mRNA sequence.
Helianthus annuus (common sunflower)
Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroidae;

REFERENCE
AUTHORS

TITLE

JOURNAL
COMMENT

FEATURES
source

1. 733
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA801"
/db_xref="taxon:4232"
/clone="QHB22C02"
/lab_host="E.coli"
/clone_lib="OH ABCDI sunflower RHA801"
/notes="Vector: pBRCNDSFIAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG TISSUE=hulls
TAG LIB=OH ABCDI sunflower RHA801
TAG_SEQ=GGTGTGCGG"

ORIGIN

Query Match 26.6%; Score 426; DB 13; Length 733;
Best Local Similarity 74.7%; Pred. No. 5.4e-59;
Matches 547; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 249 CAAGTTTATGTCGTGATGACTCTCCACTCTCAGTTTCCACCACTTCAATGGA 308
DB 733 CGAATTATGTCGTGATGACTCTCCACTCTCAGTTTCCACCACTTCAATGGA 674

QY 309 TGTGATATCTCCAAATGTTGTTACTGATCTTTCTGATGAGAGGTGGAATTTGTT 368
DB 673 TTGTGATTTCTCGAATTTGTTTACTGATCTTTCTGATGAGAGGTGGAATTTGTT 614

QY 369 TGAAGAAATGTTGAATGTTGAAGCCAGGGGTTACATTTTCTTCAGAGAAATTTGTT 428
DB 613 TGAAGAAATTTTGAATGTTGAGGTGAGGATATATATTTCTTCGAGAGTCACTGTT 554

QY 429 CCATCAATCTGGGATCAGAAACCAAGCAATCCACCACTCCAGTGAACCTAGGTT 488
DB 553 TCACAGCTGTGAGACCATAGCAAGCAAGCAATTCACACATCACTGAGACCTAGGTT 494

QY 489 CTACACTAAGGCTTCAAGAGTGTCAATTTGCAAGATGATCTGGAATCTTTATGAGCT 548
DB 493 CTACACCAAGGCTTTTAAAGAAATGTCACCACTGATGATCTGTTATCTTTATGAACT 434

QY 549 CTCCTCTACTTACTGCAATATGTTTGGAGCTTATGTCAGAAACCAAGAAACCAAGCA 608
DB 433 TAGGCTAATCTGATCGAATGCAATTCGAGCTTATGTACGCAACCAAGAAATCAATCA 374

QY 609 GATTAGTGTGTTGCGCAAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 665
DB 373 GATCTCTGTGATATGCGAGAGGTTGCGGATTTGCGAGACCGATAGGACTTCCAGGAT 314


```
Qy 666 TCTGGATCTAGCCAGTACAAAGTGTAAATAGCAATCTCGGATATGAGCGTGTATTTGGCCC 725
Db |||||
313 CTTGGATAACGTACAGTACAAAGCAATGGCATCTGGCTAGCAACGCATATTTGGACC 254
Qy 726 TGGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTCTCAATGCTGGACTT 785
Db |||||
253 AGGGTTTGTGAGCACAGGAGGATAGACACGACAAAGAAATTTTGGCAATGCTGGATCT 194
Qy 786 GAAGCTTGGCCAGAAAGTCTCGATGTTGGTGTGGAAATGGTGGAGGTGACTTTTACAT 845
Db |||||
193 GAAACCGGGGCAGAAAGTTTAGACGTTGGATGTGGAATCGGGGAGGCGCACTTCTATAT 134
Qy 846 GCGGAGACCTTTGATGTTGAGGTTGTGGATTTGATCTCTCGTTAAATATGATTTTCCTT 905
Db |||||
133 GGCTGAAAACCTTGGACGTTGATGTTGTGGCAATGATCTTCTGCAATATGATCTCTTT 74
Qy 906 TGCCCTTGAGGCTTCTATTGGGCTTAAATGTCGTTGAGTTGAGGTAGCAGATTGCAC 965
Db |||||
73 TGCTCTTGACGTGCATTTGGGCTCCAAATGTTCAATTTGAATTCGAGTTGCTGACTGTAC 14
Qy 966 CAAGATAAATA 977
Db |||||
13 CAAGAAATCATA 2
```

Search completed: August 2, 2004, 08:17:45
Job time : 4447 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 04:36:43 ; Search time 6481 Seconds
(without alignments)
10713.708 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacaccgtgtatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rdi.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	100.0	1602	6	BD084055
2	1602	100.0	1602	6	BD093370
3	1602	100.0	1779	8	AB080186
4	1138.6	71.1	2235	6	AX250551
5	1138.6	71.1	2252	8	AF237633
6	875	54.6	1794	8	AF328858
7	857.4	53.5	1870	8	AB090883
8	853.4	53.3	1656	8	AF428454
9	851.8	53.2	1459	8	AY133811
10	851.8	53.2	1660	8	AY063866
11	834	52.1	1770	8	AY058175
12	834	52.1	1809	8	AF367299
13	834	52.1	1878	8	AF197940
14	832.2	52.0	1476	8	AY091683
15	832.2	52.0	1911	8	AY093093
16	822.2	51.3	1750	8	AY139479
17	766.8	47.9	2069	8	AK069137
18	748.4	46.7	1834	8	AY136372
19	700.6	43.7	1782	8	AY065971
20	621.2	38.8	1120	6	AX250553
21	302.6	18.9	724	8	HVU234432
22	209	13.0	3221	8	AK102037
23	189.4	11.8	74968	8	AB019230
24	187.2	11.7	29508	8	AC079676
25	187.2	11.7	105937	8	AC012679
26	186.6	11.6	95295	8	AC030883
27	180.6	11.3	16886	8	AP003248
28	168.2	10.5	144305	2	AC137078
29	166.6	10.4	141715	2	AC130725
30	166.6	10.4	158911	2	AC136492
31	160.2	10.0	111754	2	AC138199
32	149.4	9.3	280	6	AR251394
33	149.4	9.3	286	6	AR252133
34	125.8	7.9	274	6	AR248084
35	109.2	6.8	260	6	AR245517
36	101	6.3	174	6	AR250330
37	100.4	6.3	274	6	AR250329
38	75.8	4.7	110000	2	PFMAL13_16
39	74.4	4.6	1220	3	AY429590
40	63.2	3.9	23614	3	U64834
41	63.2	3.9	140702	2	AC006888
42	58.4	3.6	2522	8	ATU80192
43	57	3.6	171	6	AR248566
44	56.6	3.5	3016	3	DDU11631
45	56.6	3.5	6424	3	AC116924

ALIGNMENTS

RESULT 1	BD084055	1602 bp	DNA
LOCUS	BD084055	1602 bp	DNA
DEFINITION	Environmental stress-tolerant gene.		
ACCESSION	BD084055		
VERSION	BD084055.1	GI:22629665	
KEYWORDS	JP 2001333784-A/20.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1602)		
AUTHORS	Yamada, A., Otsuki, Y., and Saito, T.		
TITLE	Environmental stress-tolerant gene		
JOURNAL	Patent: JP 2001333784-A 20 04-DEC-2001;		
	JAPAN SCIENCE AND TECHNOLOGY CORP		

COMMENT	OS	Sueda japonica
	PN	JP 2001333784-A/20
	PD	04-DEC-2001
	PF	19-JUL-2000 JP 2000219649
	PI	AKIYO YAMADA, YOSHIHIRO OZEKI, TAKEO SAITO
	PC	C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02//
	PC	C12N1/21,
	PC	C12P21/08, C12N15/00, C12N5/00
	CC	Environmental stress-tolerant gene
	PH	Location/Qualifiers
	FT	Key (1) . (1419).
	CDs	Location/Qualifiers
FEATURES	1.	1602
	source	/organism="unidentified"
ORIGIN	/mol	type="genomic DNA"
	/db_xref	="taxon:32644"
Query Match	100.0%;	Score 1602; DB 6; Length 1602;
	Best Local Similarity	100.0%; Pred. No. 0;
Matches 1602;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	CACACGTTGATTAAACCAATTGAAGCTATGATCTCGAATCTCAAGCTTCTCATCTTGAC 60
Db	1	CACACGTTGATTAAACCAATTGAAGCTATGATCTCGAATCTCAAGCTTCTCATCTTGAC 60
Qy	61	AAGAGAAACGTCCTGAGATTCTTTCAATGCTTCGGCTCTTGAAGCAAAATGCTCTTG 120
Db	61	AAGAGAAACGTCCTGAGATTCTTTCAATGCTTCGGCTCTTGAAGCAAAATGCTCTTG 120
Qy	121	GAACTTGGGGCTGGTATTGGTCGTTTACTGGTGAATTTGGCTGAGAAAGCTGGCCAGGTT 180
Db	121	GAACTTGGGGCTGGTATTGGTCGTTTACTGGTGAATTTGGCTGAGAAAGCTGGCCAGGTT 180
Qy	181	ATTGCTCTGGATTTCATTTGAGAGTCTATCAAGAGAAATCAAGTATCAATGGGCACTAC 240
Db	181	ATTGCTCTGGATTTCATTTGAGAGTCTATCAAGAGAAATCAAGTATCAATGGGCACTAC 240
Qy	241	AAAATGTCAAGTTTATGTGTGATGATCTCTCCACTCTCAGTTTCCACCACT 300
Db	241	AAAATGTCAAGTTTATGTGTGATGATCTCTCCACTCTCAGTTTCCACCACT 300
Qy	301	TCAATGATGATATCTCCAAATGGTTACTCATGTATCTTCTCATGAAGAGGTGGAA 360
Db	301	TCAATGATGATATCTCCAAATGGTTACTCATGTATCTTCTCATGAAGAGGTGGAA 360
Qy	361	AATTGGTTGAAAGATGTTGAATGGTTGAGCGCGGGTTACATTTCTTCAGAGAA 420
Db	361	AATTGGTTGAAAGATGTTGAATGGTTGAGCGCGGGTTACATTTCTTCAGAGAA 420
Qy	421	TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCACTACCTGAA 480
Db	421	TCTTGTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCACTACCTGAA 480
Qy	481	CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCAATTGCAAGATGGATCTGAAACTCT 540
Db	481	CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCAATTGCAAGATGGATCTGAAACTCT 540
Qy	541	TATGAGCTCTCCCTACTAGCTGCAAAATGATATGAGCTTATGTTCAGAAACAGAAAC 600
Db	541	TATGAGCTCTCCCTACTAGCTGCAAAATGATATGAGCTTATGTTCAGAAACAGAAAC 600
Qy	601	CAGAACAGATTAAGTTGGTTGGCAAAAGTTGATCTTAAGGATGATTAAGGGGTTCCAG 660
Db	601	CAGAACAGATTAAGTTGGTTGGCAAAAGTTGATCTTAAGGATGATTAAGGGGTTCCAG 660
Qy	661	CGATTCTGATTAAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 720
Db	661	CGATTCTGATTAAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 720
Qy	721	GGCCCTGGTTATGTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 780
Db	721	GGCCCTGGTTATGTTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 780

Qy	781	GACTTGAAGCCTGCGCAGAAAGTCTTGBATGTTGTTGGATTTGGTGGAGGTGACTTTT 840
Db	781	GACTTGAAGCCTGCGCAGAAAGTCTTGBATGTTGTTGGATTTGGTGGAGGTGACTTTT 840
Qy	841	TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGAATTTGATCTCTCCGTTAAATATGAT 900
Db	841	TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGAATTTGATCTCTCCGTTAAATATGAT 900
Qy	901	TCCTTGGCCCTTGAGCGTTCTATTTGGCTTAATGCTGTTGAGTTTGGATTTGAGTAGACAT 960
Db	901	TCCTTGGCCCTTGAGCGTTCTATTTGGCTTAATGCTGTTGAGTTTGGATTTGAGTAGACAT 960
Qy	961	TGCACCAAGATAAACTACCTGATAAATCTTTTGTATGATCATCTATAGCGGTGACACCAT 1020
Db	961	TGCACCAAGATAAACTACCTGATAAATCTTTTGTATGATCATCTATAGCGGTGACACCAT 1020
Qy	1021	CTGCATATTCAGGACAAAGCTGCGTTGTTTGTAGATCTCTACAAATGGTTGAAGCCAGCA 1080
Db	1021	CTGCATATTCAGGACAAAGCTGCGTTGTTTGTAGATCTCTACAAATGGTTGAAGCCAGCA 1080
Qy	1081	GSTAAAGTTCTAATCAGTGATCTGCAAGAAAGCTGCTCCACCTCACCTCACTGAATTCGCC 1140
Db	1081	GSTAAAGTTCTAATCAGTGATCTGCAAGAAAGCTGCTCCACCTCACCTCACTGAATTCGCC 1140
Qy	1141	GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGCAATATATGGCAGATGCTT 1200
Db	1141	GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGCAATATATGGCAGATGCTT 1200
Qy	1201	AAAGTGTGCTGATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCAATTCGAGTT 1260
Db	1201	AAAGTGTGCTGATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTCAATTCGAGTT 1260
Qy	1261	CTACGGAAGGAACTAGAGACTGTTGAGAGAAAGGATGTTGTTTATGATGATTTCTCT 1320
Db	1261	CTACGGAAGGAACTAGAGACTGTTGAGAGAAAGGATGTTGTTTATGATGATTTCTCT 1320
Qy	1321	GAGGAGGATTAACAATGACATTTGTTGAGGTTGGAATGATAAGTTGGGAGGAACTGCCAAG 1380
Db	1321	GAGGAGGATTAACAATGACATTTGTTGAGGTTGGAATGATAAGTTGGGAGGAACTGCCAAG 1380
Qy	1381	GGTGAGCAACGATGGGGTCTGTTGTTCCGAGAGAGGAGTGAAGATCAAGTTGCGGCAC 1440
Db	1381	GGTGAGCAACGATGGGGTCTGTTGTTCCGAGAGAGGAGTGAAGATCAAGTTGCGGCAC 1440
Qy	1441	GGCACTGTGCAATTTCTAGTATTAATCTTCAATGTTTTCATGTAATGATCTTCTACATGT 1500
Db	1441	GGCACTGTGCAATTTCTAGTATTAATCTTCAATGTTTTCATGTAATGATCTTCTACATGT 1500
Qy	1501	AAATTTGCCAATAAGTTCGATTCGACACTGTAAGATGATTAATCATATTTTATCTTTT 1560
Db	1501	AAATTTGCCAATAAGTTCGATTCGACACTGTAAGATGATTAATCATATTTTATCTTTT 1560
Qy	1561	AATTAATCATGATTTATGCAAAAAA 1602
Db	1561	AATTAATCATGATTTATGCAAAAAA 1602

RESULT 2
BD093370 1602 bp DNA linear PAT 27-AUG-2002
LOCUS Environmental stress-tolerant gene.
DEFINITION BD093370
ACCESSION BD093370
VERSION BD093370.1 GI:22638958
KEYWORDS WO 0106006-A/20.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1602)
Yamada, A., Ozeki, Y. and Saito, T.
Environmental stress-tolerant gene
Patent: WO 0106006-A 20 25-JAN-2001.
JOURNAL JAPAN SCIENCE AND TECHNOLOGY CORP, AKIYO YAMADA, YOSHIHIRO OZEKI,

TAKEO SAITO
 OS Sueada japonica
 PN WO 0106006-A/20
 PD 25-JAN-2001
 PR 19-JUL-2000 WO 2000JP004862
 AKIYO YAMADA, YOSHIHIRO OZeki, TAKEO SAITO
 PC C1201/68, C12N15/29, C07K14/415, C07K16/16, C12P21/02, A01H5/00 CC
 FH Key Location/Qualifiers
 FT CDS (1). (1419).
 FEATURES
 source Location/Qualifiers
 1..1602
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 1602; DB 6; Length 1602;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CACACGGTTGATTTAAACCAATGAAGCTATGATGCTCCGATCTCAAGCTTCTGATCTTGAC 60
 DB 1 CACACGGTTGATTTAAACCAATGAAGCTATGATGCTCCGATCTCAAGCTTCTGATCTTGAC 60
 QY 61 AAAGAAGAACGTCCTGAGATCTTTCAATGCTCCGCTCTTGAAGAAAATGCCCTCTTG 120
 DB 61 AAAGAAGAACGTCCTGAGATCTTTCAATGCTCCGCTCTTGAAGAAAATGCCCTCTTG 120
 QY 121 GAACTTGGGGCTGGTATTTGGTCTGTTTAACTGGTGAATTTGGCTGAGAAAGCTGCCAGGTT 180
 DB 121 GAACTTGGGGCTGGTATTTGGTCTGTTTAACTGGTGAATTTGGCTGAGAAAGCTGCCAGGTT 180
 QY 181 ATTGCTCTGGATTTCAATGAGAGTGCTATCAAGAAGAAATGAATCAATCAATGGCACTAC 240
 DB 181 ATTGCTCTGGATTTCAATGAGAGTGCTATCAAGAAGAAATGAATCAATCAATGGCACTAC 240
 QY 241 AAAATGTCMAAGTTATGTCGTGATGTCGATCTTCCCACTCTCAATTTCCCAACACAT 300
 DB 241 AAAATGTCMAAGTTATGTCGTGATGTCGATCTTCCCACTCTCAATTTCCCAACACAT 300
 QY 301 TCATTTGGATGTGATATCTCCAAATTTGGTGTACTCATGTATCTTCTCATGCAAGAGGTGGA 360
 DB 301 TCATTTGGATGTGATATCTCCAAATTTGGTGTACTCATGTATCTTCTCATGCAAGAGGTGGA 360
 QY 361 AATTGGTTGAAAGAAATGTTGAATGTTCAAGCCAGGGGTTACATTTCTTCAGAGAA 420
 DB 361 AATTGGTTGAAAGAAATGTTGAATGTTCAAGCCAGGGGTTACATTTCTTCAGAGAA 420
 QY 421 TCTTGTTTCCATCAATCTGGGATCAAAAGCAAGCAATCCCACTACCGTGAA 480
 DB 421 TCTTGTTTCCATCAATCTGGGATCAAAAGCAAGCAATCCCACTACCGTGAA 480
 QY 481 CCTAGGTTCTACACTAAGGCTCTCAAGAGGTGATTTTCAGAGTGGATCTGGAACCTCT 540
 DB 481 CCTAGGTTCTACACTAAGGCTCTCAAGAGGTGATTTTCAGAGTGGATCTGGAACCTCT 540
 QY 541 TATGAGCTCTCCCTACTAGCTGCAAAATGATTTGGAGCTTATGTCAGAAACAAAGAAAC 600
 DB 541 TATGAGCTCTCCCTACTAGCTGCAAAATGATTTGGAGCTTATGTCAGAAACAAAGAAAC 600
 QY 601 CAGAACCAAGATAGTTGGTTGGTGGCAAAAAGTTGATTTAAGATGATTAAGGGGTTCCAG 660
 DB 601 CAGAACCAAGATAGTTGGTTGGTGGCAAAAAGTTGATTTAAGATGATTAAGGGGTTCCAG 660
 QY 661 CGATTTCTGGATCTAGCCAGTACAGTGTAAATAGCAATCTTCGGATATGAGGGGTATTT 720
 DB 661 CGATTTCTGGATCTAGCCAGTACAGTGTAAATAGCAATCTTCGGATATGAGGGGTATTT 720
 QY 721 GGCCTCTGGTTATGTTAGCACTGCGAGATATCAAAACCAAGAGGTTGTGTCAATGCTG 780
 DB 721 GGCCTCTGGTTATGTTAGCACTGCGAGATATCAAAACCAAGAGGTTGTGTCAATGCTG 780

QY 781 GACTTGAGAGCTTGGCCAGACGGTCTCTGGATGTTGGTTGTGGAATTTGGTGGAGTGACTTT 840
 DB 781 GACTTGAGAGCTTGGCCAGACGGTCTCTGGATGTTGGTTGTGGAATTTGGTGGAGTGACTTT 840
 QY 841 TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT 900
 DB 841 TACATGGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT 900
 QY 901 TCTTTGGCCCTTGAGCGTCTTATTGGGCTTAATGCTGCTGCTGAGTTGAGGCTAGCAGAT 960
 DB 901 TCTTTGGCCCTTGAGCGTCTTATTGGGCTTAATGCTGCTGCTGAGTTGAGGCTAGCAGAT 960
 QY 961 TGCACCAAGATTAACCTACCTGATTAATCTTTTGAATGTCATCTATAGCCGTGACACATT 1020
 DB 961 TGCACCAAGATTAACCTACCTGATTAATCTTTTGAATGTCATCTATAGCCGTGACACATT 1020
 QY 1021 CTGCATATTTCAGGACCAAGCCTGCTGTTTAGATCTCTTCTACAAATGGTTGAAGCCAGCA 1080
 DB 1021 CTGCATATTTCAGGACCAAGCCTGCTGTTTAGATCTCTTCTACAAATGGTTGAAGCCAGCA 1080
 QY 1081 GGTAAAGTTCTTAATCAGTGATTACTGCAAGAAAGCTGGTCCACCTCACCTGAATTCGCC 1140
 DB 1081 GGTAAAGTTCTTAATCAGTGATTACTGCAAGAAAGCTGGTCCACCTCACCTGAATTCGCC 1140
 QY 1141 GCTTACATTAAAGCAGAGGGGATATGATCTCCATGATGTAAAGGAATAATGGGCAGATGCTT 1200
 DB 1141 GCTTACATTAAAGCAGAGGGGATATGATCTCCATGATGTAAAGGAATAATGGGCAGATGCTT 1200
 QY 1201 AAAAGATGCTGATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTTCAATCGAGTT 1260
 DB 1201 AAAAGATGCTGATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTTCAATCGAGTT 1260
 QY 1261 CTACGGAAGAACTAGAGACTGTTGAGAGAAAGGATGTGTTCAATAGTATGTTCTCT 1320
 DB 1261 CTACGGAAGAACTAGAGACTGTTGAGAGAAAGGATGTGTTCAATAGTATGTTCTCT 1320
 QY 1321 GAGGAGGATTAACAATGACATTTGTTGGAGGTTGGAAATGATAAGTTGCGAGGACTGCCAAG 1380
 DB 1321 GAGGAGGATTAACAATGACATTTGTTGGAGGTTGGAAATGATAAGTTGCGAGGACTGCCAAG 1380
 QY 1381 GGTGAGCAACCTGGGCTGTTGGTGGCGAGGATAGAACTGAGCAGTTTCAATCGAGTT 1440
 DB 1381 GGTGAGCAACCTGGGCTGTTGGTGGCGAGGATAGAACTGAGCAGTTTCAATCGAGTT 1440
 QY 1441 GGCACCTGTCGATTTCTTCTAGTATTAATCTCAATGTTTTCATGTAATGTAATTTCTACATGT 1500
 DB 1441 GGCACCTGTCGATTTCTTCTAGTATTAATCTCAATGTTTTCATGTAATGTAATTTCTACATGT 1500
 QY 1501 AAAATTTGCCAATAAGTTGCAATTTGCGAGACTGTAAAGATGATTAATCAATTTTATCTTTT 1560
 DB 1501 AAAATTTGCCAATAAGTTGCAATTTGCGAGACTGTAAAGATGATTAATCAATTTTATCTTTT 1560
 QY 1561 AATTAATCATCGATTTATGCAAAAAAATAAAAAAATAAAAAA 1602
 DB 1561 AATTAATCATCGATTTATGCAAAAAAATAAAAAAATAAAAAA 1602

RESULT 3

AB080186
 LOCUS
 DEFINITION
 Sueda japonica PEAMT mRNA for phosphoethanolamine
 N-methyltransferase, complete cds.
 ACCESSION
 AB080186
 VERSION
 AB080186.1 GI:28436073
 KEYWORDS
 Sueda japonica
 Sueda japonica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Amaranthaceae; Sueda.
 ORGANISM
 Sueda japonica
 Sueda japonica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Caryophyllales; Amaranthaceae; Sueda.
 REFERENCE
 1 Yamada, A., Nozawa, G.T., Tanimoto, S. and Ozeki, Y.
 Glycinebetaine synthesis in Sueda japonica

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1779)
AUTHORS Yamada, A., Nozawa, G.T. and Ozeki, Y.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) Akiyo Yamada, Tokyo University of
Agriculture and Technology, Department of Biotechnology; Naka-cho
2-24-16, Koganei, Tokyo 184-8588, Japan
(E-mail: yamadenec.tuat.ac.jp, Tel.81-423-88-7239,
Fax.81-423-88-7239)
FEATURES
Location/Qualifiers
1..1779
/organism="Suaeda japonica"
/mol_type="mRNA"
/db_xref="taxon:90346"
/country="Japan: Saga, Ariake Sea"
1..1779
/gene="PEAMT"
115..1599
/gene="PEAMT"
/codon_start=1
/product="phosphoethanolamine N-methyltransferase"
/protein_id="BAC57432.1"
/db_xref="GI:28436074"
/translation="MAASGMELMQBRDVPKVVHVDLTIEAMWLDQSADLDKKE
RPEILSMPLPECHGLELGGI(GRTGELAEKAGOVIALDFIESAIKKNVINGHYK
NVKFMCAVTSFTUSFEPSLDVIFSNWLLWYLSDEVENLVERMLKXPGGVIFER
ESCFHQSDHKKRKNPHYREPRPTKAFKCHLQDGSNTELSLCKICGAYEN
KKNQHSIWLKQKVDSDKDFQFLDTQYKCNLSILRYRVPFGYVSTGGYITKE
FVSLDLKPGOKVLDVCGGIGGGDFYMAETFDVEVPDLSVNMISFALERSILGKCA
VEFEVADCTKINYPDNSFDVYSRDTLHIDQKPLFRSFKMLKPGGVLSIDYCK
AGPPSPFPAAYIKQGYDLHVKYQMLKADGFVDVLAEDRETFPIRLAKELTVE
KEKQVFIQSDPSEEDYNDIVGMDNKLRLRTAKGRWGLFVANKK"

source
gene
CDS
ORIGIN
Query Match 100.0%; Score 1602; DB 8; Length 1779;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CACACGGTGTGATTAACCAATGAAGCTATGATGCTCGAATCTCAAGCTTCTGATCTTGAC 60
Db 178 CACACGGTGTGATTAACCAATGAAGCTATGATGCTCGAATCTCAAGCTTCTGATCTTGAC 237
Oy 61 AAGAAGAACGCTCTGAGATCTTTCAATGCTCCGCTCTTGAAGGAAATGCCCTCTTG 120
Db 238 AAGAAGAACGCTCTGAGATCTTTCAATGCTCCGCTCTTGAAGGAAATGCCCTCTTG 297
Oy 121 GAATCTGGGCTGGTATGCTGTTTACTGGTGAATGGCTGAGAAAGCTGGCCAGGTT 180
Db 298 GAATCTGGGCTGGTATGCTGTTTACTGGTGAATGGCTGAGAAAGCTGGCCAGGTT 357
Oy 181 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAAATGAAGTAATCAATGGGCACTAC 240
Db 358 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAAATGAAGTAATCAATGGGCACTAC 417
Oy 241 AAAAAATCAAGTTTATGCTGATGCTGATGCTCTCCACTCTCAGTTTCCACCAACAT 300
Db 418 AAAAAATCAAGTTTATGCTGATGCTGATGCTCTCCACTCTCAGTTTCCACCAACAT 477
Oy 301 TCATGGAATGATATCTCCAAATGGTTACTCATGATCTTTCTGATGAAGAGGTGGAA 360
Db 478 TCATGGAATGATATCTCCAAATGGTTACTCATGATCTTTCTGATGAAGAGGTGGAA 537
Oy 361 AATTGGTTCGAAGATGTTGAATGGTTGAAGCCAGGGGTTCATTTCTTCAGAGAA 420
Db 538 AATTGGTTCGAAGATGTTGAATGGTTGAAGCCAGGGGTTCATTTCTTCAGAGAA 597
Oy 421 TCTTGTTCATCATCTGGGATCACAAACCAAGCAATCCACCCACTACCGTGA 480
Db 598 TCTTGTTCATCATCTGGGATCACAAACCAAGCAATCCACCCACTACCGTGA 657
Oy 481 CTAAGGTTCTACACTAAGGCTTTCAAGAGTGTCTATTTGCAAGATGATCTGGAACCT 540
Db 658 CTAAGGTTCTACACTAAGGCTTTCAAGAGTGTCTATTTGCAAGATGATCTGGAACCT 717

541 TATGAGCTCTCCCTACTAGCTGCAAAATGTATTTGGAGCTTATGTGAGAAACAGAAAAAC 600
Db 718 TATGAGCTCTCCCTACTAGCTGCAAAATGTATTTGGAGCTTATGTGAGAAACAGAAAAAC 777
Oy 601 CAGAACCAAGATAGTTGGTGTGGCAAAAGTTGATTTCTTAAGGATGATAAGGGGTTCAG 660
Db 778 CAGAACCAAGATAGTTGGTGTGGCAAAAGTTGATTTCTTAAGGATGATAAGGGGTTCAG 837
Oy 661 CGATTTCTGATACTAGCCAGTACAAAGTGTATAGCAATTTCTGCGATATGAGCGTGTATT 720
Db 838 CGATTTCTGATACTAGCCAGTACAAAGTGTATAGCAATTTCTGCGATATGAGCGTGTATT 897
Oy 721 GGCCTGTGTTATGTTAGCACTGAGAGATATGAACCAACCAAGAGTGTGTCATGCTG 780
Db 898 GGCCTGTGTTATGTTAGCACTGAGAGATATGAACCAACCAAGAGTGTGTCATGCTG 957
Oy 781 GACTTGAAGCTCTGCGAGAGGTCCTGGATGTTGGTGTGGATTTGCTGGAGGTGACTTT 840
Db 958 GACTTGAAGCTCTGCGAGAGGTCCTGGATGTTGGTGTGGATTTGCTGGAGGTGACTTT 1017
Oy 841 TACATGCGGAGACCTTTGATGTTGAGTTGTTGGATTTGATCTCTCCGTTAATGATTT 900
Db 1018 TACATGCGGAGACCTTTGATGTTGAGTTGTTGGATTTGATCTCTCCGTTAATGATTT 1077
Oy 901 TCCTTTGGCTTGGAGCGTTCTATTGGGCTTAAATGTGCTTGGTGTGAGTTTGGAGTAGCAGAT 960
Db 1078 TCCTTTGGCTTGGAGCGTTCTATTGGGCTTAAATGTGCTTGGTGTGAGTTTGGAGTAGCAGAT 1137
Oy 961 TGCACCAAGATAAATACCTGATTAACCTTTTGAATGTCATCTATAGCGCTGAGACCAATT 1020
Db 1138 TGCACCAAGATAAATACCTGATTAACCTTTTGAATGTCATCTATAGCGCTGAGACCAATT 1197
Oy 1021 CTGCATATTCAGGACCAAGCTCTGCTTCTTGTAGATCTCTTACAAATGGTTGAGCCAGCA 1080
Db 1198 CTGCATATTCAGGACCAAGCTCTGCTTCTTGTAGATCTCTTACAAATGGTTGAGCCAGCA 1257
Oy 1081 GGTAAAGTTCTAATCAGTATTAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1140
Db 1258 GGTAAAGTTCTAATCAGTATTAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1317
Oy 1141 GCTTACATTAAGCAGAGGGGATGATCTCATGATGATAAGGAATATGGGCGAGTCTT 1200
Db 1318 GCTTACATTAAGCAGAGGGGATGATCTCATGATGATAAGGAATATGGGCGAGTCTT 1377
Oy 1201 AAGATGCTGGAATTTGATGTTCTTGGAGGATAGAACTGAGCAGTTCATTTCCAGTT 1260
Db 1378 AAGATGCTGGAATTTGATGTTCTTGGAGGATAGAACTGAGCAGTTCATTTCCAGTT 1437
Oy 1261 CTACGAGGAACTAGAGACTGTTGAGAGGAAAGGATGCTGTTCAATGATGATTTCTCT 1320
Db 1438 CTACGAGGAACTAGAGACTGTTGAGAGGAAAGGATGCTGTTCAATGATGATTTCTCT 1497
Oy 1321 GAGGAGGATTAACATGACATTTGTTGGAGTTGGAATCATTAAGTTGCGGAGGACTGCGAAG 1380
Db 1498 GAGGAGGATTAACATGACATTTGTTGGAGTTGGAATCATTAAGTTGCGGAGGACTGCGAAG 1557
Oy 1381 GGTGACCAAGATGGGCTCTGTTCTGCGCAAGAGAGTGAAGAACTCAGTTCCGCACT 1440
Db 1558 GGTGACCAAGATGGGCTCTGTTCTGCGCAAGAGAGTGAAGAACTCAGTTCCGCACT 1617
Oy 1441 GGCACTGTGCGATTTCTAGTATTAATCTTCAATGTTTTCATGATGATCTTCTACATGT 1500
Db 1618 GGCACTGTGCGATTTCTAGTATTAATCTTCAATGTTTTCATGATGATCTTCTACATGT 1677
Oy 1501 AAAATTTGCCAATAAGTTGATTTTCCGACACTGTAAGATGATTAATCATATTTTATCTTTT 1560
Db 1678 AAAATTTGCCAATAAGTTGATTTTCCGACACTGTAAGATGATTAATCATATTTTATCTTTT 1737
Oy 1561 AATTAATCATGATTTTATGCAAAAAAATAAAAAAATAAAAAA 1602
Db 1738 AATTAATCATGATTTTATGCAAAAAAATAAAAAAATAAAAAA 1779

VERSION	AF237633.1	GI:7407188	Db
KEYWORDS	Spinacia oleracea (spinach)		
SOURCE	Spinacia oleracea		Qy
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Amaranthaceae; Spinacia.		Db
REFERENCE	1. (bases 1 to 2252)		Qy
AUTHORS	Nuccio, M.L., Ziemak, M.J., Henry, S.A., Weretilnyk, E.A. and Hanson, A.D.		Db
TITLE	cDNA cloning of phosphoethanolamine N-methyltransferase from spinach by complementation in <i>Schizosaccharomyces pombe</i> and characterization of the recombinant enzyme		Qy
JOURNAL	J. Biol. Chem. 275 (19), 14095-14101 (2000)		Db
MEDLINE	20361526		Qy
PUBMED	10799484		Db
REFERENCE	2. (bases 1 to 2252)		Db
AUTHORS	Nuccio, M.L. and Hanson, A.D.		Qy
TITLE	Direct Submission		Db
JOURNAL	Submitted (21-FEB-2000) Horticultural Sciences, University of Florida, P.O. Box 110690, Gainesville, FL 32611, USA		Db
FEATURES	Location/Qualifiers		
source	1..2252		Qy
gene	/organism="Spinacia oleracea"		Db
CDS	/mol_type="mRNA"		Qy
	/db_xref="taxon:3562"		Db
	/tissue_type="salinized leaf"		Qy
	1..2252		Db
	/gene="PRMT"		Qy
	254..1738		Db
	/gene="PRMT"		Qy
	/EC number="2.1.1.103"		Db
	/function="catalyzes S-adenosylmethionine dependent N-methylation of phosphoethanolamine, phosphonomethylmethanolamine and phosphodimethylmethanolamine"		Qy
	/codon_start=1		Db
	/product="phosphoethanolamine N-methyltransferase"		Qy
	/protein_id="AA661950.1"		Db
	/db_xref="GI:7407188"		Qy
	/translation="MAASANGVLQREVEFKYWI EHSVDLTVEAMMLDSQASDLDKVE RPEVLSMLPYEGKSLVLEGLGFRFTGELAKSQAIVLADFI ESVIKKNS INGHYK NPKWPCADVTSPGSLNIPNSVDIIFSNWLLTSLDEEVLVERLWKLKPGGIYFPR ESHFQSGDHRKSNPHTYRPRFYTKI FKECHMODDSNSYELSLIGCKTGAYVKS KKNQNOISLWQKVDSDDDKGPORFLDSQYKFNISLRVRFVGGYVSTGGLLETKS FVSKLDLPKQKVLQVCGGIGGGDFYMAENYDVEVVGIDLSINMLSPALERSIGLKA VEFVADCTKQKVPENSFDVIYSRDTILHIOKPALFRSFHKLKPGKVLISDYCKS AGTSASFPAYIRGORGDLHDVAYGKMLKDGAFVBYAENRTDQPIQVLKELDALE QEKDQDFDDFSEEDINDIVDGMKAKLVRTTEGGQWGLFIAKKM"		Db
ORIGIN			
Query Match	71.1%; Score 1138.6; DB 8; Length 2252;		
Best Local Similarity	82.6%; Pred. No. 4.9e-270;		
Matches 1328; Conservative	0; Mismatches 274; Indels 5; Gaps 2;		
Qy	1	CACACGGTGTATTAACCATTAAGCTATGATGCTGATCTCAAGCTTCTGATCTTGAC	60
Db	317	CACTCTGTGTATTAAGCTATGATGCTGATCTCAAGCTTCTGATCTTGAC	376
Qy	61	AAAGAAAGCGCTGAGATCTTTTCAATGCTTCGGCTCTTGAAGAAATATGCTCTTG	120
Db	377	AAAGTGAAGCGACCTGAGTACTTTCCATGCTTCCACCTTATGAAGGAAGTCTGTCTTA	436
Qy	121	GAATCTGGGCTGGTATGCTGCTTTTACTGGTGAATTTGGCTGAGAAAGCTGGCCAGGTT	180
Db	437	GAATCTGGTCTGGTATGCTGCTTTTACTGGTGAATTTGGCCGAGAAAGCTAGCCAGTC	496
Qy	181	ATTGCTCTGATTTCAATGAGTCTATCAAGAGAAATCAAGTATCAATGAGTCTAC	240
Db	497	ATCGCTCTGATTTCAATGAGTCTATCAAGAGAAATCAAGTATCAATGAGTCTAC	556
Qy	241	AAAAATGTCAGTTTATGCTGATGATGCTCTCCCACTCTCAGTTTCCCAACAT	300
Db	1637	GAGGAGGATTTATAACGACATAGTTGATGGTTGGGAAGCCGANGTTGGTGAAGCTACAGAG	1696
Db	557	AAAAATGTCAGTTTATGCTGATGATGATGCTCCAGTCTCAAGTCTCAACATTTCCACCAAT	616
Qy	301	TCATTTGATGATGATTTCTCCAAATGGTTTACTCATGTATCTTTCTGTGATGAAGGTGAA	360
Db	617	TCCGTGATATCATATCTCCAAATGGTACTCATGTATCTTTCTGTGATGAAGGTGAG	676
Qy	361	AATTTGGTTCAGAAATGTTGAAATGGTTGAAGCAGGGGGTTACATTTCTTTCAGAGAA	420
Db	677	CGTCTGGTTGAAGGATGTTGAATGGTTGAAGCCAGGAGGATACATTTCTTTCAGAGAA	736
Qy	421	TCTTGTTCATCATCTGCGGATTCAGAAACCAAGCAATCCACCCTTACCTGAA	480
Db	737	TCATTGTTTTCATCAATCAGGATCAAGCCGCAAGCAATCCAAACCTTACCTGAA	796
Qy	481	CCTAGGTTTCTACATAAGGCTTTCAAGAGTGTCAATTTGCAAGATGATCTCTGAAAATCT	540
Db	797	CCTAGGTTTCTACATAAGGCTTTCAAGAGTGTCAATTTGCAAGATGATCTCTGAAAATCT	856
Qy	541	TATGAGCTCTCCCTTACTAGCTGCAATCTATTTGGAGCTTATGTCAGAAACAGAAAC	600
Db	857	TATGAGCTCTCCCTTACTAGCTGCAATCTATTTGGAGCTTATGTCAGAAACAGAAAC	916
Qy	601	CAGAACAGATTAAGTTGGTTGTCGCAAAAGTTGATCTTAAGATGATAAGGGTTCAG	660
Db	917	CAGAACAGATTAAGTTGGTTGTCGCAAAAGTTGATCTTAAGATGATAAGGGTTCAG	976
Qy	661	CGATTTCTGGATCTAGCCAGTACAAGTGTAAATAGCATCTCGCATATAGAGGCTGATTT	720
Db	977	CGATTTCTGGATCTAGCCAGTACAAGTGTAAATAGCATCTCGCATATAGAGGCTGATTT	1036
Qy	721	GGCCCTGGTTATGTTAGCACTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG	780
Db	1037	GGTCCTGGTTATGTTAGTACCGAGGATCGAAACCAACCAAGAGTTTGTGTCAATGCTG	1096
Qy	781	GACTTTGAAGCTCGCCAGAAAGTCTCGGATGTTGGTGTGGAAATTTGGTGGAGTCACTTT	840
Db	1097	GACTTTGAAGCTCGCCAGAAAGTCTTAGATGTGGTGTGGTGTGGTGTGGAGTCACTTT	1156
Qy	841	TACATGGCGAGACCTTTGATGTTGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	900
Db	1157	TACATGGCGAGACCTTTGATGTTGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	1216
Qy	901	TCCTTTGCCCTTGAGCGTTCTATTGGGCTTAAATGCTGTTGAGTGTGGTGTGGTGTGGT	960
Db	1217	TCCTTTGCCCTTGAGCGTTCTATTGGGCTTAAATGCTGTTGAGTGTGGTGTGGTGTGGT	1276
Qy	961	TGCAACAAGATAAACTACCTGATACTCTTTGATGTCATCTATAGCCGTGACACCATTT	1020
Db	1277	TGCAACAAGATAAACTACCTGATACTCTTTGATGTCATCTATAGCCGTGACACCATTT	1336
Qy	1021	CTGCAATTTACGACAGCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	1080
Db	1337	CTGCAATTTACGACAGCCCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	1396
Qy	1081	GGTAAAGTTCTTAATCAGTGTACTGCAAGAAAGCTGGTCCACCTTCACTGAAATTCGCC	1140
Db	1397	GGTAAAGTTCTTAATCAGTGTACTGCAAGAAAGCTGGTCCACCTTCACTGAAATTCGCC	1456
Qy	1141	GCTTACATTAAGCAGAGGGATGATGCTCCATGATGTAAGGAATATGGCAGAGTCTT	1200
Db	1457	GCTTACATTAAGCAGAGGGATGATGCTCCATGATGTAAGGAATATGGCAGAGTCTT	1516
Qy	1201	AAAGATGCTGGATTTGATGTTTGGCCAGAGATAGAACTGAGCAGTTCATTCAGGTT	1260
Db	1517	AAAGATGCTGGATTTGATGTTTGGCCAGAGATAGAACTGAGCAGTTCATTCAGGTT	1576
Qy	1261	CTACCGAAGGAACTAGAGACTGTTGAGAGGAAAGGATGTTTCAATGATGATTTCTCT	1320
Db	1577	CTACCGAAGGAACTAGAGACTGTTGAGAGGAAAGGATGTTTCAATGATGATTTCTCT	1636
Qy	1321	GAGGAGGATTTATTAACGACATAGTTGATGGTTGGGAAGCCGANGTTGGTGAAGCTACAGAG	1380

QY	429	CCATCAATCTGGGATCACAAACGCAAAAGCAATCCACCACTACCGTGAACCTAGCTT	488
DB	453	CCATCAATCTGGGACGACGCAAGCAAGCAAAACCCACTACCTAGCGTGAACCCAGATT	512
QY	489	CTACATCAAGCTTCAAGAGTGCTTTTTCAGAGTGCATGCTGCAAGTCTTATGAGCT	548
DB	513	CTACATCAAGCTTTCAGAGTGCATGCTGCAAGTCTTTCAGAGTGCATGCTTATGAGCT	572
QY	549	CTCCTACTAGCTGCAAAATGTTATGAGCTTATGTCAGAAACCAAGAAACCAAGACCA	608
DB	573	CTCTATGCTGGCTGCAAAATGCTATGAGCTTATGTCAGAAACCAAGAAACCAAGACCA	632
QY	609	GATTAGTGGTGGGCAAAAGCTTCTTAAAGATGATGAGGCTTTCAGAGCTTCTT	668
DB	633	GATTAGTGGTGGGCAAAAGCTTCTTAAAGATGATGAGGCTTTCAGAGCTTCTT	692
QY	669	GGATCTAGGCTGATGAGTGTATGAGCTTCTGAGTATGAGCTTATGAGCTTCTT	728
DB	693	GGATCTAGGCTGATGAGTGTATGAGCTTCTGAGTATGAGCTTATGAGCTTCTT	752
QY	729	TTATGTTAGCTGAGGATGATGAAACCAACCAAGAGTGTGTCAATGCTGGCTGAA	788
DB	753	ATATGTTAGCTGAGGATGATGAAACCAACCAAGAGTGTGTCAATGCTGGCTGAA	812
QY	789	CGCTGGGCGAGGCTGAGT	848
DB	813	ACCGGACGAGGCTGAGT	872
QY	849	GGAGACCTTTGATGTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	908
DB	873	TGAGAAATTCGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATG	932
QY	909	CTTGAGGCTTCTATGCTGCTTAAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	968
DB	933	ACTGAGGCGGCTGAGT	992
QY	969	GATTAACCTACCTGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1028
DB	993	CAAAACATATCCGATGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1052
QY	1029	TCAGGACGAGCTGCTGTTGATGATCTTCTCAAAATGTTGAAAGCCAGGAGGTAAAGT	1088
DB	1053	CCAGACGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1112
QY	1089	TCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1148
DB	1113	TCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1172
QY	1149	TAAGACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1208
DB	1173	AAACACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1232
QY	1209	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1268
DB	1233	AGGCTTTGACGAGCTTATGCTGAGGACCGTACTGATGATGATGATGATGATGATGAT	1292
QY	1269	GGATCTAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1328
DB	1293	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1352
QY	1329	TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1388
DB	1353	TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1412
QY	1389	ACGTTGGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1423
DB	1413	GAATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1447

RESULT 9
 AV133811
 LOCUS
 1459 bp mRNA linear PLN 18-SEP-2002

DEFINITION	Arabidopsis thaliana clone U11999 putative phosphoethanolamine N-methyltransferase (At1g48600) mRNA, complete cds.
ACCESSION	AV133811
VERSION	AV133811.1
KEYWORDS	GI:22136801
SOURCE	FLI CDNA
ORGANISM	Arabidopsis thaliana (thale cress)
REFERENCE	Arabidopsis thaliana Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1459) Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A. Arabidopsis Open Reading Frame (ORF) Clones 2 (bases 1 to 1459) Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE	Direct Submission
JOURNAL	Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
REFERENCE	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL CDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
COMMENT	The Salk, Stanford, PGE (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
FEATURES	Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs. Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank. Location/Qualifiers 1..1459 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="1" /clone="U11999" /ecotype="Columbia" /note="This clone is in pUNI 51." 1..1459 /gene="At1g48600" 1..1428 /gene="At1g48600" /codon_start=1 /evidence="experimental" /product="putative phosphoethanolamine N-methyltransferase" /protein_id="AA091745.1" /db_xref="GI:22136802" /translation="MEHSSDUTVEAMMLDSKASDLKRPVLSLIPPEGKSVLEL"

QY 1149 TAAGCAGAGGATATGATCTCCATGATGTAAAGGAATATGGCAGATCTTAAAGATGC 1208
 Db |||||
 QY 1189 AAAACAAGAGGATATGATCTACATGATGTTCAAGCTTACGACAGATCTGAAGAGGC 1248
 Db |||||
 QY 1209 TGGATTGTGATGTTCTGCGAGGATAGACAGTGGAGTTCATTCAGATCTTACGGAA 1268
 Db |||||
 QY 1249 AGGCTTTGACGACGTTATCGTGAAGACCGTACTGATCTGTTGTAACAAGTCTCAGGCG 1308
 Db |||||
 QY 1269 GGAATCTAGAGACTGTTGAGAGGAAAAGGATGTTTCTATTAGTATCTCTGAGGAGGA 1328
 Db |||||
 QY 1309 TGAATTAGAAAAGTGGAGAAAGAAAGGAAGAAATTCATCAGCGACTTCTCAGAGAGGA 1368
 Db |||||
 QY 1329 TTACAATGACATTTGTTGACGCTTGGCAATGATAGTTGCGGAGGACTGCCAAGGTTGACA 1388
 Db |||||
 QY 1369 TTACAATGACATTTGTTGACGCTTGGCAATGATAGTTGCGGAGGACTGCCAAGGTTGACA 1428
 Db |||||
 QY 1389 ACGATGGGCTGTTGCTGTTGCGGAGGAGGAGTGA 1423
 Db |||||
 QY 1429 GAAATGGGATTTATTCATAGCCGACGAAGTAA 1463
 Db |||||

RESULT 11

AY058175
 LOCUS
 DEFINITION
 Arabidopsis thaliana AT3g18000/MEB5_22 mRNA, complete cds.
 AY058175
 VERSION
 FLJ_CDNAs.
 Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 1770)
 Cheuk R., Chen H., Kim C.J., Koesena E., Meyers M.C., Banh J.,
 Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M.,
 Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K.,
 Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
 Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K.,
 Davis R.W., Theologis A. and Ecker J.R.
 Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 1770)

Cheuk R., Chen H., Kim C.J., Koesena E., Meyers M.C., Banh J.,
 Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M.,
 Narusaka M., Nguyen M., Onodera C.S., Palm C.J., Pham P.K.,
 Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
 Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S., Shinozaki K.,
 Davis R.W., Theologis A. and Ecker J.R.
 Direct Submission

TITLE

Submitted (28-AUG-2001) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

COMMENT

Riken Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Cheuk, R., Chen, H.,
 Kim, C.J., Koesena, E., Meyers, M.C., Shinn, P., Banh, J., Bower, L.,
 Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Sakurai, T.,
 Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
 Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES

source

1..1770
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clones="RAPL09-10-N16 (R14028)"
 /ecotype="Columbia"
 1..215

5'UTR

216..1691

/note="putative methyltransferase"

/codon_start=1

/product="AT3g18000/MEB5_22"

/protein_id="AAL25589.1"

/db_xref="GI:1664879"

/translation="MAASYEERDIOKNYWIHSADLTVEAMMLDSRADLDKEERPE

VLSLPVEGKSVLELGAGIGRTGELAKAGELIALQFIDNVIKENSINGHYNVK

FMCAVTSPLDKITDGLDIFSNMLLWLSKEVELLAERWGVKIGTIFPHSC

PHQSGSKRKNPTHYREPFPYKVFQRCQTRDAAGNSFELSMIGCKICGIVKRNKN

QNQICWIKQVSSNDRGFORPLDNVQYKSSGLIRYERFVGQGVSTGLSTTKEPVE

EVADCTTKRYPDMSFDVIYSDRDLHIQDKPALFRTPFKLPGKGVLSIDYCSPT

PSAESEYIKRGYDLHDVQAGQMLKDGFTDIAEDRTDQFMQVLKRELDREVEKE

EKFISDFSKEDYDDIVGGWKSLEKRCASDEQKWLFIANKN"

1692..1770

3'UTR

ORIGIN

Query Match 52.1%; Score 834; DB 8; Length 1770;
 Best Local Similarity 74.1%; Pred. No. 7.4e-35;
 Matches 1056; Conservative 0; Mismatches 370; Indels 0; Gaps 0;

QY 1 CACACGTTGATTTAAACATGAAGCTATGATCTCGATTTCTCAAGCTTCTGATCTTGCAC 60
 Db |||||
 QY 270 CATTCGCTGATCTGACTGTTTGAAGCTATGATCTTGAAGCTTCTGATCTTGCAC 329
 Db |||||
 QY 61 AAAGAAGAAAGCTCTGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAATACCTCTTG 120
 Db |||||
 QY 330 AAGGAAGAAAGCTCTGAGGACTCTCTTCTGCTCCATATGAGGCAATCACTGTG 389
 Db |||||
 QY 121 GAATCTGGGCTGATTTGCTGCTTTTCTGCTGATTTGCTGAGGAAATGCTGAGGAGTCC 180
 Db |||||
 QY 390 GAATCTGGGCTGATTTGCTGCTTTTCTGCTGATTTGCTGAGGAAATGCTGAGGAGTCC 449
 Db |||||
 QY 181 ATGCTCTGGATTTCTATGAGAGTGTCTATCAAGAAGATGAAGTAAATCAATGGGCACTAC 240
 Db |||||
 QY 450 ATTGCTCTTGAATTTCTATGATTAACGTTTATCAAGAAGATGAAGTAAATCAATGGGCACTAC 509
 Db |||||
 QY 241 AAAAATGTCAAGTTTATGTGCTGTGATGTGACTTCTCCCACTCTCAGTTTCCCAACAT 300
 Db |||||
 QY 510 AAGATGTCAAGTTTATGTGCTGTGATGTGACTTCCCTGACCTCAGATCATGATGGA 569
 Db |||||
 QY 301 TCATTTGATGTGATTTCTCAATGCTTACTCATGATTTCTTCTGATGAAGAGTGGAA 360
 Db |||||
 QY 570 TCTCTGATTTGATTTTCTCCAACTGGCTGCTCATGTATCTTCTGACAAAGAGGTGGAG 629
 Db |||||
 QY 361 AATTGGTTGAAGAAATGTTGAATGTTTGAAGCAGAGGGGTTTACATTTTCTTCAGAGAA 420
 Db |||||
 QY 630 CTTTGGCAGAAAGATGGTGGTGGATTCAGATTTGAGGAGTATCAATTTTCTTCGTGA 689
 Db |||||
 QY 421 TCTGTTTCCATCATCTGGGATCAAAACGCAAGCAATCCCACTACTACCTGGA 480
 Db |||||
 QY 690 TCTTGTCTCCCACTCAGGGACAGTAAGCGGAAATCCCACTACTACTACCTGGA 749
 Db |||||
 QY 481 CTTAGGTTCTACACTTAAGGCTTCAAGAGTGTCTATTTGCAAGATGATCTGGAATCTCT 540
 Db |||||
 QY 750 CCGGTTCTTCTTCTCAAGTCTTCTCAAGAGTGTGAGTCTGCTGCTGGAATCTCA 809
 Db |||||
 QY 541 TATGAGTCTCCCTACTTACTGCTCAATGATTTGAGGCTTATGTCAGAAACAGAAAGAAC 600
 Db |||||
 QY 810 TTTGAGTCTCTCTATGATCGGATCAAGTGTGAGGCTTATGTCAGAAACAGAAAGAAC 869
 Db |||||

QY 601 CAGAACAGATTAGTCTGTTGTCGCAAAAGTTGATTTCTAAGGATGATAAGGGGTTCCAG 660
 DB 870 CAGATCAGATTGTTGATATGCGCAGAGGTCAGCTCAGAAATGACAGAGGCTTCCAA 929
 QY 661 GATTTCTGGATACACGAGTACAGTGTATAGCATTTCTCGATATGAGGTTGATTT 720
 DB 930 GATTTCTGGATACACGAGTACAGTGTATAGCATTTCTCGATATGAGGTTGATTT 989
 QY 721 GGCCTGCTGATGTTAGTACCTGGAGGATGATGAACACCAACAGAGGTTGTTGTCATGCTG 780
 DB 990 GGCCTGCTGATGTTAGTACCTGGAGGATGATGAACACCAACAGAGGTTGTTGTCATGCTG 1049
 QY 781 GATTTCTGGATACACGAGTACAGTGTATAGCATTTCTCGATATGAGGTTGATTT 840
 DB 1050 GATTTCTGGATACACGAGTACAGTGTATAGCATTTCTCGATATGAGGTTGATTT 1109
 QY 841 TACATGCGGAGACCTTTGATGTTGAGGTTGTTGATGTTGATGTTGATGTTGATGTT 900
 DB 1110 TACATGCGGAGACCTTTGATGTTGAGGTTGTTGATGTTGATGTTGATGTTGATGTT 1169
 QY 901 TCCCTTGGCTGAGGCTTTCTATGGGCTTAAATGTTGCTGTTGAGTTGAGGTAGCAGAT 960
 DB 1170 TCCCTTGGCTGAGGCTTTCTATGGGCTTAAATGTTGCTGTTGAGTTGAGGTAGCAGAT 1229
 QY 961 TCCACAGATATACCTGATACCTTTTGTGATGCTATCTATAGCGGTCGACACCAT 1020
 DB 1230 TCCACAGATATACCTGATACCTTTTGTGATGCTATCTATAGCGGTCGACACCAT 1289
 QY 1021 TCCGATATTCAGGACAGGCTGCTGTTGTTAGATCTCTTACAAATGGTTGAGGACGAGA 1080
 DB 1290 TCCGATATTCAGGACAGGCTGCTGTTGTTAGATCTCTTACAAATGGTTGAGGACGAGA 1349
 QY 1081 GGTAAAGTCTTAATCAGTGTATCTGCAAGAAAGCTGCTCCACCTCCACCTGAAATTCGCC 1140
 DB 1350 GGTAAAGTCTTAATCAGTGTATCTGCAAGAAAGCTGCTCCACCTCCACCTGAAATTCGCC 1409
 QY 1141 GCTTACATTAAGCAGAGGAGATGATCTCATGATGATGATGATGATGATGATGATGAT 1200
 DB 1410 GATGATACATTAAGCAGAGGAGATGATCTCATGATGATGATGATGATGATGATGATGAT 1469
 QY 1201 AAGATGCTGAGTATGTTGATGTTCTTCCGAGGATGATGATGATGATGATGATGATGATGAT 1260
 DB 1470 AAGATGCTGAGTATGTTGATGTTCTTCCGAGGATGATGATGATGATGATGATGATGATGAT 1529
 QY 1261 CTACGAGGAGTACGAGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 DB 1530 CTACGAGGAGTACGAGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1589
 QY 1321 GAGGAGGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 DB 1590 GAGGAGGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1649
 QY 1381 GGTGAGCAACGAGGCTGCTGTTGCTGCTGCAAGAAAGGATGAAAGAA 1426
 DB 1650 GGTGAGCAACGAGGCTGCTGTTGCTGCTGCAAGAAAGGATGAAAGAA 1695

RESULT 12
 AF367299
 LOCUS Arabidopsis thaliana 1809 bp mRNA linear PLN 07-FEB-2002
 DEFINITION Arabidopsis thaliana AT3g18000/MB85_22 mRNA, complete cds.
 ACCESSION AF367299
 VERSION AF367299.1 GI:13605800
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 rosids; euroids; J1. Brassicales; Brassicaceae; Arabidopsi;
 1 (bases 1 to 1809)
 REFERENCE Chouk.R., Chen.H., Kim.C.J., Meyers.M.C., Shinn.P., Banh,J.,
 Bowser.L., Carninci.P., Chung.M.K., Goldsmith.A.D., Hayashizaki.Y.,

Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
 Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Arabidopsis cDNA clones
 Unpublished
 2 (bases 1 to 1809)
 Chouk.R., Chen.H., Kim, C.J., Meyers, M.C., Shinn, P., Banh, J.,
 Bowser, L., Carninci, P., Chung, M.K., Goldsmith, A.D., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Narusaka, M.,
 Nguyen, M., Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Sakurai, T.,
 Satou, M., Seki, M., Southwick, A., Toriumi, M., Yamada, K., Yu, G.,
 Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.
 Direct Submission
 Submitted (31-MAR-2001) Salk Institute Genomic Analysis Laboratory
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

COMMENT
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PCRC (SSP) Consortium members carried out the
 sequencing and annotation of the RAPL cDNAs: Shinn, P., Chen, H.,
 Chouk, R., Kim, C.J., Koesema, E., Meyers, M.C., Tracy, S.E., Banh, J.,
 Bowser, L., Chung, M.K., Goldsmith, A.D., Jones, T., Karlin-Neumann, G.,
 Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M.,
 Palm, C.J., Pham, P.K., Quach, H.L., Sakano, H., Southwick, A.,
 Tang, C.C., Toriumi, M., Yamada, K., Yu, G., Davis, R.W., Theologis, A.,
 and Ecker, J.R.

Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES
 source
 1..1809
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="RAPL09-11-124 (R12246)"
 /ecotype="Columbia"
 1..140
 141..1616
 /note="putative methyltransferase"
 /codon_start=1
 /product="AT3G18000/MB85_22"
 /protein_id="AAK32886.1"
 /db_xref="GI:13605801"
 /translation="MAASVEERDIOKNYIEHSDLTVEAMMLDSRASDLDEKPERPE
 VLSILPYEGKSLVGLGAGIGRGTGELAQAGELIALDFIDNVKKNESINGHYNNK
 FMCADVTSPDLKIDTGLSLDILFNSMLIYLSDEKVELLAERVMGVIKVGVIYFPFESK
 PHOGSKRKNSTHYRPFYKVFQBCOTDAAGNSFELSMIGCKICGAVYKNNK
 CNOICWIKVSSNDRGFORFLDNYQYKSSGILRYLVERVGGQFVSTGLETKEFVE
 RNKLKPCQKVLNCGCGIGGGDPFMAKEDVHVYVIGLISVMISPALERIGLSCSVEF
 EVADCTTHYFNSFDVITNSDILHFDKAPALFFTFKMLKPGKVLISD/CRSPT
 PSASEFYIKRGYLDHVQYGMKADGTFDVIADRDITQPMQVILKELDRVSEKE
 EKPTSDFSKEDYDDIVGWSKSLERCAESDEKQKGLFIANNK"
 1617..1809

3'UTR
 CDS

ORIGIN

Query Match 52.1%; Score 834; DB 8; Length 1809;
 Best Local Similarity 74.1%; Pred. No. 7.4e-195;
 Matches 1056; Conservative 0; Mismatches 370; Indels 0; Gaps 0;

QY 1 CAGACCGTTCATTTACCATTCAGCTATGATGCTCATCTCAAGCTTCGATCTGAC 60
 DB 195 CATTCCGCTGATCTGACTGTTGAAGCTATGATGCTGACTGAGAGCTTCGATCTCGAC 254

QNQCWTQKVSSENRGRFRLDNVOYKSGILRYRVPQGFVFGGLTETKFEVE
 KMLKPQKVLVDVCGGIGGDFYMAEKFDVHVVDLSDVNMISFALRAIGLSCSVF
 EVADCTKHYPDNSFVITYSDRTILHIQDKPALFTFFMLKPGQKVLISDYCRSPK
 PSAPSEYIKORGYDLHDVQAYQKLDAGFTDVIADERTDQFMQVLKRELDVREKEK
 EKTFSDFSKEDYDDIVGGWKSLEKRCASDEQKGLFIANKQ"
 361..660
 /gene="NM1"
 /note="Region: SAM-binding domain"
 361..387
 /gene="NM1"
 /note="Region: Motif I"
 427..438
 /gene="NM1"
 /note="Region: post-I"
 544..570
 /gene="NM1"
 /note="Region: Motif II"
 631..660
 /gene="NM1"
 /note="Region: Motif III"
 1048..1341
 /gene="NM1"
 /note="Region: SAM-binding domain"
 1048..1074
 /gene="NM1"
 /note="Region: Motif I"
 1114..1128
 /gene="NM1"
 /note="Region: post-I"
 1231..1254
 /gene="NM1"
 /note="Region: Motif II"
 1312..1341
 /gene="NM1"
 /note="Region: Motif III"
 52.18; Score 834; DB 8; Length 1878;
 Best Local Similarity 74.18; Pred. No. 7.4e-195;
 Matches 1056; Conservative 0; Mismatches 370; Indels 0; Gaps 0;
 ORIGIN
 1 CACACGCTGGATTAAACCATGAGCTATGCTGCTGATCTCAAGCTCTGATCTTGAC 60
 256 CATTCCGCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAC 315
 61 AAAGAGAACTCTGAGATCTTTCAATGCTTCCGCTCTTTGAAGGAAATGCTCTTG 120
 316 AAGGAAGAACTCTGAGGACTCTCTTGTCTCTCCATATGAAGGCAATCAGTGTG 375
 121 GAACCTGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 376 GAACCTGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 435
 181 ATTGCTCTGATTTCAATGAGTGTATCAAGAGAAATGAAGTAAATCAATGGGCACTAC 240
 436 ATTGCTCTGATTTCAATGAGTGTATCAAGAGAAATGAAGTAAATCAATGGGCACTAC 495
 241 AAAAATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 496 AAGAAATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
 301 TCATGAGTGTATTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 556 TCTCTGATGCTGATTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
 361 AATTGCTGTAAGATTTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 616 CTTTGGGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
 421 TCTGTTTCAATCAATCTGGGATCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAG 480
 676 TCTGCTTCCACCAATCAGGGACAGTAAAGCGAAATCCACCCCTCACTACCGTGA 735

481 CTTAGGTTCTACACTAAGGCTTCAAGAGTGTCTATTTGCAAGATGCTATCGAACTCT 540
 736 CCGCGTTCTATTTCAAGGCTTTTCAAGAGTGTCTATTTGCAAGATGCTATCGAACTCT 795
 541 TATGAGCTCTCCCTACTTATGCTCAAAATGTTATGAGGCTTATGTCAGAAACAAAGAAAC 600
 796 TTTGAGCTCTCTATGATCGGATGCAATGCTATTTGAGGCTTATGTCAGAAACAAAGAAAT 855
 601 CAGAACCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 856 CAGAAATCAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 915
 661 CGATTTCTGGAATCTAGCCAGTACAAAGTGTAAATAGCAATCTCTGCGATATGAGGCTGTATT 720
 916 GCTTTCTTGGCAATGCTCAATATCAAAATCCAGTGGAAATCTAGCTATGAGGCTGTCTTT 975
 721 GGCCTGCTGATTTAGCACTGAGGATATGAACCAACCAAGAGCTTTGCTGCTGCTGCT 780
 976 GGCCAAGGCTTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035
 781 GACTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 1036 AATCTGAACCAAGCAAGAAAGTCTTAGATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1095
 841 TACATGCGGAGACTTTGATGTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCT 900
 1096 TACATGCTGAGGCTTTGATGTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCTTTGAGGCT 1155
 901 TCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 1156 TCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1215
 961 TGCACCAAGATAAACTACCTGATACTCTTTTGTATGCTCTATAGCGCTGACACCACTTT 1020
 1216 TGCACCAACCAACACTACCCAGATAATTCGTTGATGCTATTTACAGCGCTGACACTATT 1275
 1021 CTGCATATTCAGGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 1276 CTGCATATTCAGGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335
 1081 GGTAAAGTCTTAATCAGTGTATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 1336 GGTAAAGTCTTAATCAGTGTATTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1395
 1141 GCTTACATTAAGCAGAGGCTGATGCTCTCCATGATGTAAGGAAATATAGGCAAGTGTCTT 1200
 1396 GAGTACATCAACAGAGGCTGATGCTCTCCATGATGTAAGGAAATATAGGCAAGTGTCTT 1455
 1201 AAAGATGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1456 AAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
 1261 CTACGGAAGGAACTAGAGACTGTTGAGAGGAAAGGATGCTGCTGCTGCTGCTGCTGCTGCT 1320
 1516 CTGAAAAGTGAATAGACAGGCTGAGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 1575
 1321 GAGGAGATTAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 1576 AAAGAGATTAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1635
 1381 GGTGAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1426
 1636 GATGAGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1681

RESULT 14
 AV091683
 LOCUS
 DEFINITION
 Arabidopsis thaliana AT3g18000/ME85_22 mRNA, complete cds.
 AV091683
 ACCESSION
 AV091683.1
 VERSION
 FLI CDNA
 KEYWORDS
 Arabidopsis thaliana (thale cress)


```

1075 DB CTGCACATCCAGCAACACGAGCTGTTAGGACTTCTTCAATGGCTTAAACCGGGA 1134
1081 QY GGTAAAGTCTTAATCAGTGTACTGCAAGAAGCTGGTCCACCCCTCACTGAATTCGCC 1140
1135 DB GGTAAAGTCTTCAATCAGGACTACTGTAGAAGCCCCAAAACCTCCATCTGCTGAGTTTCA 1194
1141 QY GCTTACATTAGCAGGAGGATATGATCTCCATGATGTAAGCATATGCGCATGCTT 1200
1195 DB GAGTACATCAACAGAGAGATATGATCTCCATGATGTAAGCATATGCGCATGCTT 1254
1201 QY AAGAGTCTGATTTGTTGATGTTCTGCGAGAGATAGAACTGACAGTTCATTCGAGTT 1260
1255 DB AAAGAGCTGCTGCTCACTGATGATCGCAGAGAGACCTGATGATGATGATGATGATG 1314
1261 QY CTACGAGGAGAACTAGAGACTGTTGAGAGAGAAAGGATGCTGCTCATGATGATGCTT 1320
1315 DB CTGAAAGCTGATGATGAGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1374
1321 QY GAGGAGGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
1375 DB AAGAGGATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1434
1381 QY GGTGAGCAAGAGAGGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTT 1422
1435 DB GATGAGCAGAAATGGGAGCTTTTCATGCGCCACACAGAAATTA 1476

```

```

RESULT 15
LOCUS AY093093 1911 bp mRNA linear PLN 21-APR-2002
DEFINITION Arabidopsis thaliana unknown protein (At3g17990:At3g18000) mRNA,
complete cds.
ACCESSION AY093093
VERSION AY093093.1 GI:20260387
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.
1 (bases 1 to 1911)
Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M.,
Palm, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H.,
Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,
Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K.,
Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
Submitted (27-MAR-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu

```

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PSEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Southwick, A., Nguyen, M., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Southwick, A., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
1..1911
/organism="Arabidopsis thaliana"

```

gene
CDS
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="3"
/clone="RAPL08-10-M03 (R17629)"
/ecotype="Columbia"
/notes="This clone is in pBluescript"
1..1911
/gene="At3g17990:At3g18000"
216..1691
/gene="At3g17990:At3g18000"
/notes="This cDNA spans two annotated genes At3g17990 and At3g18000"
/codon_start=1
/product="unknown protein"
/protein_id="AA13092.1"
/db_xref="GI:20260388"
/translation="MAASVEERDIQKWIHESADLTVEAMLDSPASDLKEERPE
VLGILPRYSGKSVLELGAIGSTGELAGKAGELIALDPIQNVIKONESINCHYKVK
FMADVTSDPKLITDGLDILFBNWLLATLSDKEVELLAERKVKWIKGVYIPRESC
PHUGDSKRKSNFTHREPRFYKVFQECYTRDAAGNSFELSMIGCKCIAYIKPKN
QNOICWIMQVSSNDURGFQRPIDNVQYKSGILRYKRVFGGFPVSTGGLTKEPVE
KMLKPGQKVLDDGCGIGGGDFYMAEKFDVHVVGIDLSVNMISFALERAIGLSCSPK
EVADCTTKHVPDPSFDVYSRDTILHIQDKPALPFFKMLKPGKVLISDYCSKPK
PSAEFSEYIKQRGYDLHDVQAYGOMLKDAGFTDVAEDRTDQFMQVLKRELDREKEX
EKFISDFSEYDIDYVGGWKSLEKRCASDEQKWLFIANKN"

```

ORIGIN

```

Query Match      52.0%  Score 832.4  DB 8; Length 1911;
Best Local Similarity 74.0%  Pred. No. 1.8e-194;
Matches 1055; Conservative 0; Mismatches 371; Indels 0; Gaps 0;

QY 1 CACACCGTTGATTAACCATGGAAGTATGATGCTGATCTCAAGCTTCTGATCTTGCAC 60
DB 270 CATTCCCGCTGATCTGACTGTTGAAGCTATGATGCTTGACTCGAGAGCTTCTGATCTGCAC 329
QY 61 AAGAAGAAAGCTCTGAGATTTCTTCAATGCTTCCGCTCTTGAAGGAAATGCTCTTGG 120
DB 330 AAGGAGAAAGCTCTGAGGACTCTCTTGTCTCCCTCATATGAAGGCAATCACTGTTG 389
QY 121 GAATCTGGGCTGATGATGCTGTTTACTGCTGATGATGCTGCTGAGAAAGCTGCGCAGGTT 180
DB 390 GAATCTGGGCTGATGATGCTGTTTACTGCTGATGATGCTGCTGAGAAATGCTGCTGCTG 449
QY 181 ATTGCTCTGGAATTTCTGAGAGTGTCTATCAAGAAGAAATGAAGTAAATCAATGGGCACTAC 240
DB 450 ATTGCTCTGGAATTTCTGAGAGTGTCTATCAAGAAGAAATGAAGTAAATCAATGGGCACTAC 509
QY 241 AAAAATCTCAAGTTTATGCTGCTGATGCTGCTTCCCTCTGCTGCTGCTGCTGCTGCTGCT 300
DB 510 AAAAATCTCAAGTTTATGCTGCTGATGCTGCTTCCCTCTGCTGCTGCTGCTGCTGCTGCT 569
QY 301 TCATTGAGTGTATGATTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 570 TCATTGAGTGTATGATTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629
QY 361 AATTGGTTGAAAGAAATGTTGAATGCTTGAAGCAGAGGGGGTTTACATTTTCTTCAGAGAA 420
DB 630 CTTTGGCAGAAAGGAGTGTGCTGCTGATCAAGTTTGGAGGATACATTTTCTTCGTTGAA 689
QY 421 TCTGTTTTCATCATCTGGGATCACAAAGCAAGCAATCCACCCTACCTGCTGCTGCTGCTGCT 480
DB 690 TCTGTTTTCATCATCTGGGATCACAAAGCAAGCAATCCACCCTACCTGCTGCTGCTGCTGCT 749
QY 481 CTTAGGTTTCTACACTAAGGCTTTCAAGAGTGTCTATTTGCAAGATGAGTCTGGAAGAACTCT 540
DB 750 CTTAGGTTTCTATTTCAAGGCTTTTCAAGAGTGTCTGAGACTCGGAGTGTCTGCTGCTGCTG 809
QY 541 TATGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB 810 TTTGAGTCTCTATGATCGGATCGAATGCAAGTGCATTTGAGGCTTATGTCACAGACAGAGAA 869
QY 601 CAGAACCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

```

Db 870 CAGAAATCAGATTGTTGGATATGCGCAGAGGTGAGCTCAGAAATGACAGAGGCTTCCAA 929
Qy 661 CGAATTTCTGGATACAGCAGTACAGGTGTAATAGCAATCTCGATATGAGCGTGTATTT 720
Db 930 CGTTTCTTGGCAATGTCCAAATCAAAATCCAGTGGATCTAGGCTATGAGCGTGTCTTT 989
Qy 721 GGGCTGGTATGTTAGCAGTGGAGGATATGAACCAACCAAGAGTTTGTGCAATGCTG 780
Db 990 GGGCAAGGTTTGTGAGCAGTGGTGGACTTGAGCAACCAAGAAATTTGTGAGAAATG 1049
Qy 781 GACTTGAAGCTGGCCAGAGGCTCTCGATGTTGGTGTGGAAATGGTGGAGTGCATTT 840
Db 1050 AATCTGAACACCAAGCAGACAGAAAGTCTTAGATGTTGGGTGGCATGGTGGAGGTGACTTC 1109
Qy 841 TACATGGCGGAGACCTTTGATGTTGAGTGTGGATTTGATCTCTCCGTTAATATGAT 900
Db 1110 TACATGGCTGAGAGTTGATGTTGACGTTTGTGGTATGATCTTTCTGTCACATGATC 1169
Qy 901 TCCTTGGCTTGAGCGTTCTATTTGGGCTTAAATGTGCTGTGAGTTGAGTAGCAGAT 960
Db 1170 TCTTTGGCATTTGAAACGCTGCTATTTGGACTCAGCTGCTCGGTTGGGTTGAGGTTGCTGAT 1229
Qy 961 TGCACCAAGATAAACTACCTGATAACTCTTTTGTGATCATCTATAGCCGTCACACCAAT 1020
Db 1230 TGCACCAAAACACTACCCAGATAATTCGTTGATGTCATTTACAGCCGTCACACTAT 1289
Qy 1021 CTGCATATTCAGGACAGCCCTGCTGTTTGTAGTCTTCTCAAAATGGTTGAAGCCAGGA 1080
Db 1290 CTGCATCTCAAGACAAACAGCCCTGTTTGTAGACTTTCTTCAAAATGGCTTAAACCGGA 1349
Qy 1081 GGTAAAGTTCTAATCAGTGATTAAGTCAAGAAAGCTGCTCCACCTCAGCTGAATTCGCC 1140
Db 1350 GGTAAAGTTCTCATCGCGACTACTGTAGAAGCCCAAACTCCATCTGCTGATTTCA 1409
Qy 1141 GCTTACATTAAGCAGAGGGATATGATCTCCATGATGTAAAGGAATATGGCAGATGCT 1200
Db 1410 GAGTACATCAACAGAGAGGATATGATCTCCTGACGTTTACAGCTTATGACAGATGCTA 1469
Qy 1201 AAGATGCTGGATTTGTTGATGTTCTTGGCGAGGATAGAACTGAGCAGTTCAATCGAGTT 1260
Db 1470 AAGAGCGTGGCTTCACTGATGTTGATCGCAGAGGACCGTACTGATCAGTTTATGCAAGTC 1529
Qy 1261 CTACGGAAGAACTAGAGACTGTTGAGAAAGAAAGGATGTTCAATAGTATTTCTCT 1320
Db 1530 CTGAAACGTGAATTAGACAGGCTGGAGAAAGAAAGAAATTCATCTCCGACTTCTCC 1589
Qy 1321 GAGGAGTTTCAATGACATTTGTTGGAGGTTGGATGATAGTTGGAGGACTGCCAAG 1380
Db 1590 AAGAGGATTACGATGACATTTGTTGGAGGATGGAAAGTCAAAGCTGGAGAGGTTGATCG 1649
Qy 1381 GGTGAGCAACGATGGGCTCTGCTGTTGCCAAGAAAGTGAAGAA 1426
Db 1650 GATGAGCAAAATGGGACTTTTCATCGCCCAACAGAAATTAAGCAA 1695

Search completed: August 2, 2004, 07:03:12
Job time : 6486 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 08:20:03 ; Search time 6481 Seconds
(without alignments)
10713.708 Million cell updates/sec

Title: US-10-031-331b-39
Perfect score: 1602
Sequence: 1 cacacgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sta.*

28: em_un.*

29: em_vi.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pln.*

35: em_hg_rtd.*

36: em_hg_nam.*

37: em_hg_vrt.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1602	100.0	1602	6	BD084055	BD084055 Environme
2	1602	100.0	1602	6	BD093370	BD093370 Environme
3	1602	100.0	1779	8	AB080186	AB080186 Suada ja
4	35	2.2	1459	8	AV133811	AV133811 Arabidops
5	35	2.2	1656	8	AP428454	AP428454 Arabidops
6	35	2.2	1660	8	AY063866	AY063866 Arabidops
7	35	2.2	95235	8	AC020889	AC020889 Genomic e
8	32	2.0	1120	6	AX250553	AX250553 Sequence
9	32	2.0	2235	6	AX250551	AX250551 Sequence
10	32	2.0	2252	8	AF237633	AF237633 Spinacia
11	31	1.9	1347	5	BC049504	BC049504 Danio rer
12	31	1.9	138116	2	BX324004	BX324004 Danio rer
13	31	1.9	168179	2	BX293550	BX293550 Danio rer
14	31	1.9	172587	5	BX005125	BX005125 Zebrafish
15	31	1.9	175745	2	BX530060	BX530060 Danio rer
16	31	1.9	184574	5	BX294666	BX294666 Zebrafish
17	31	1.9	187570	2	BX649276	BX649276 Danio rer
18	31	1.9	226787	2	BX004886	BX004886 Danio rer
19	31	1.9	247330	2	BX530409	BX530409 Danio rer
20	31	1.9	251370	5	AL929533	AL929533 Zebrafish
21	30	1.9	1875	9	BC053645	BC053645 Homo sapi
22	30	1.9	2600	9	BC055010	BC055010 Homo sapi
23	30	1.9	105045	9	AC117494	AC117494 Homo sapi
24	30	1.9	129782	5	BX120005	BX120005 Zebrafish
25	30	1.9	180305	5	BX004768	BX004768 Zebrafish
26	30	1.9	197549	5	AL935335	AL935335 Zebrafish
27	30	1.9	347050	3	PFA929351	PFA929351 Plasmodiu
28	29	1.8	786	3	AB080223	AB080223 Samia cyn
29	29	1.8	1016	8	AY349618	AY349618 Arabidops
30	29	1.8	1102	9	BC016308	BC016308 Homo sapi
31	29	1.8	1429	8	AF508150	AF508150 Rheum tat
32	29	1.8	1449	9	BC005331	BC005331 Homo sapi
33	29	1.8	1734	9	AK000538	AK000538 Homo sapi
34	29	1.8	1833	9	BC000954	BC000954 Homo sapi
35	29	1.8	1870	8	AB090883	AB090883 Aster tri
36	29	1.8	1891	9	HSM807589	HSM807589 Homo sapi
37	29	1.8	2742	9	HSM801378	HSM801378 Homo sapi
38	29	1.8	3518	9	BC007647	BC007647 Homo sapi
39	29	1.8	6514	5	BC063203	BC063203 Silurana
40	29	1.8	61793	2	AC087703	AC087703 Homo sapi
41	29	1.8	71284	8	F21B23	F21B23 Arabidops
42	29	1.8	79389	10	AL607127	AL607127 Mouse DNA
43	29	1.8	83141	9	AL592067	AL592067 Human DNA
44	29	1.8	85702	8	AC020665	AC020665 Arabidops
45	29	1.8	110000	2	PFMAU13_02	Continuation (3 of

ALIGNMENTS

RESULT 1	LOCUS	BD084055	1602 bp	DNA	linear	PAT 27-AUG-2002
BD084055	DEFINITION	Environmental stress-tolerant gene.				
BD084055	ACCESSION	BD084055.1	GI:22629665			
BD084055.1	VERSION	JP 2001333784-A/20.				
BD084055.1	KEYWORDS	unidentified				
BD084055.1	SOURCE	unclassified.				
BD084055.1	ORGANISM	unclassified.				
BD084055	REFERENCE	1 (bases 1 to 1602)				
BD084055	AUTHORS	Yamada,A., Ozeki,Y. and Saito,T.				
BD084055	TITLE	Environmental stress-tolerant gene				
BD084055	JOURNAL	Patent: JP 2001333784-A 20 04-DEC-2001;				
BD084055		JAPAN SCIENCE AND TECHNOLOGY CORP				

COMMENT	OS	Sueada japonica
PD	JP 2001333784-A/20	
PP	04-DEC-2001	
PI	19-JUL-2000 JP 2000219649	
PC	AKIYO YAMADA, YOSHIHIRO OZEKI, TAKBO SAITO	
PC	C12N15/09, A01H5/00, C07K14/415, C07K16/16, C12N5/10, C12P21/02//	
PC	C12N1/21, C12N15/00, C12N5/00	
PC	C12P21/08, C12N15/00, C12N5/00	
CC	Environmental stress-tolerant gene	
FH	Key Location/Qualifiers	
FT	CDS (1)..(1419).	
FEATURES	Location/Qualifiers	
source	1..1602	
	/organism="unidentified"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32644"	
ORIGIN		
Query Match	100.0%; Score 1602; DB 6; Length 1602;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1602; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	CACACCGTTGATTAACCAATGAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60
Db	1	CACACCGTTGATTAACCAATGAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60
Qy	61	AAAGAGAGCGTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG 120
Db	61	AAAGAGAGCGTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG 120
Qy	121	GAACTTGGGCTGATTTGGTTCGTTTACTGTTGATGATGCTGAGAAAGCTGGCGAGTT 180
Db	121	GAACTTGGGCTGATTTGGTTCGTTTACTGTTGATGATGCTGAGAAAGCTGGCGAGTT 180
Qy	181	ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAATCAATGAGGCACTAC 240
Db	181	ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAATCAATGAGGCACTAC 240
Qy	241	AAATGTCAGTTTATGTCGTATGCTGATGCTTCCACCTCTCAGTTTCCACCAAT 300
Db	241	AAATGTCAGTTTATGTCGTATGCTGATGCTTCCACCTCTCAGTTTCCACCAAT 300
Qy	301	TCATTGATGATATCTCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGAA 360
Db	301	TCATTGATGATATCTCCAAATGGTTACTCATGTATCTTTCTGATGAAGAGGTGAA 360
Qy	361	AAATTGGTTCAGAGATGTTGAATGCTTGAAGCAGGCGGTATCATTTCTTCAGAGAA 420
Db	361	AAATTGGTTCAGAGATGTTGAATGCTTGAAGCAGGCGGTATCATTTCTTCAGAGAA 420
Qy	421	TCATTGTTCCATCAATCTGGGATCAAAACGCAAAAGCAATCCACCTACCGTGAA 480
Db	421	TCATTGTTCCATCAATCTGGGATCAAAACGCAAAAGCAATCCACCTACCGTGAA 480
Qy	481	CTTAGGTTCTACACTAAGGCTTCAAAAGAGTGTCAATTTGCAAGATGATCTGAAACTCT 540
Db	481	CTTAGGTTCTACACTAAGGCTTCAAAAGAGTGTCAATTTGCAAGATGATCTGAAACTCT 540
Qy	541	TATGAGCTCTCCCTACTAGCTGCAATGATATGAGCTTATGTCAGAAACAGAAAC 600
Db	541	TATGAGCTCTCCCTACTAGCTGCAATGATATGAGCTTATGTCAGAAACAGAAAC 600
Qy	601	CAGAACCGAGTATGTTGTTGTCGCAAAAGTGTATTTCAAGATGATGAGGGTTCAG 660
Db	601	CAGAACCGAGTATGTTGTTGTCGCAAAAGTGTATTTCAAGATGATGAGGGTTCAG 660
Qy	661	CGATTTCTGGATCTAGCCAGTCAAGTGTATGATGATTTCTCGGATATGAGCGTATTT 720
Db	661	CGATTTCTGGATCTAGCCAGTCAAGTGTATGATGATTTCTCGGATATGAGCGTATTT 720
Qy	721	GGCCCTGTTATGTTAGCACTGGAGGATATGAAACCAACCAAGAGTTTGTGTCAATGCTG 780
Db	721	GGCCCTGTTATGTTAGCACTGGAGGATATGAAACCAACCAAGAGTTTGTGTCAATGCTG 780

Qy	781	GACTTGAAGCCTGCGCAAGAGGTCCTGGATGTTGGTTGGGAATTTGGTGGAGGTGACTTT 840
Db	781	GACTTGAAGCCTGCGCAAGAGGTCCTGGATGTTGGTTGGGAATTTGGTGGAGGTGACTTT 840
Qy	841	TACATGGCGGAGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT 900
Db	841	TACATGGCGGAGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT 900
Qy	901	TGCTTTGGCTTGGAGGCTTCTATTGGCTTAAATGCTGTTGAGTTTCAAGGTAGCAGAT 960
Db	901	TGCTTTGGCTTGGAGGCTTCTATTGGCTTAAATGCTGTTGAGTTTCAAGGTAGCAGAT 960
Qy	961	TGCACCAAGATAAACTACCTGATAACTCTTTTGAATGCTATATAGCGTGCACCATTT 1020
Db	961	TGCACCAAGATAAACTACCTGATAACTCTTTTGAATGCTATATAGCGTGCACCATTT 1020
Qy	1021	CTGCATATTTACGACCAAGCCTGCGTTGTTTAGATCTCTTACAAATGGTTGAAGCCAGCA 1080
Db	1021	CTGCATATTTACGACCAAGCCTGCGTTGTTTAGATCTCTTACAAATGGTTGAAGCCAGCA 1080
Qy	1081	GSTAAAGTTCTAATCAGTGATTTACTGCAAGAAAGCTGGTCCACCTCAGCTGAATTCGCC 1140
Db	1081	GSTAAAGTTCTAATCAGTGATTTACTGCAAGAAAGCTGGTCCACCTCAGCTGAATTCGCC 1140
Qy	1141	GCTTACATTAAGCAGAGGGGATGATCTCCATGATGTAAGGAATATGGGCGAGATGCTT 1200
Db	1141	GCTTACATTAAGCAGAGGGGATGATCTCCATGATGTAAGGAATATGGGCGAGATGCTT 1200
Qy	1201	AAAGATGCTGGAATTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTTCATTCGAGTT 1260
Db	1201	AAAGATGCTGGAATTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTTCATTCGAGTT 1260
Qy	1261	CTACGAGAAACTAGAGACTGTTGAGAAAGGAATGTTTCATTTAGTATTTCTCT 1320
Db	1261	CTACGAGAAACTAGAGACTGTTGAGAAAGGAATGTTTCATTTAGTATTTCTCT 1320
Qy	1321	GAGAGGATTTACAATGACATTTGAGAGTTGCAATGATATGATTTGCGAGGAGCTGCCAAG 1380
Db	1321	GAGAGGATTTACAATGACATTTGAGAGTTGCAATGATATGATTTGCGAGGAGCTGCCAAG 1380
Qy	1381	GGTGAGCAAGGATGGGCTGTTGCTGCGAAGAGAGAGTGAAGAATCAGTTCCGCACT 1440
Db	1381	GGTGAGCAAGGATGGGCTGTTGCTGCGAAGAGAGAGTGAAGAATCAGTTCCGCACT 1440
Qy	1441	GGCACTGTCGATTTCTAGTATTAATTTCAATGTTTTCATGTAATGATCTTACATGT 1500
Db	1441	GGCACTGTCGATTTCTAGTATTAATTTCAATGTTTTCATGTAATGATCTTACATGT 1500
Qy	1501	AAATTTGCCAATGATTTGCAATTTCCAGACTGTAAGATGATTAAATCATATTTTATCTTTT 1560
Db	1501	AAATTTGCCAATGATTTGCAATTTCCAGACTGTAAGATGATTAAATCATATTTTATCTTTT 1560
Qy	1561	AATTAATCATGATTTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602
Db	1561	AATTAATCATGATTTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602

RESULT 2	
BD093370	1602 bp DNA linear PAT 27-AUG-2002
LOCUS	Environmental stress-tolerant gene.
DEFINITION	BD093370.1
ACCESSION	BD093370.1 GI:22638958
VERSION	WO 0106006-A/20.
KEYWORDS	unidentified
SOURCE	unclassified.
ORGANISM	1 (bases 1 to 1602)
REFERENCE	Yamada, A., Ozeki, Y. and Saito, T.
AUTHORS	Environmental stress-tolerant gene
TITLE	Patent: WO 0106006-A 20 25-JAN-2001;
JOURNAL	JAPAN SCIENCE AND TECHNOLOGY CORP, AKIYO YAMADA, YOSHIHIRO OZEKI,

COMMENT
 OS Sueoda japonica
 PN WO 0106006-A/20
 PD 25-JAN-2001
 PF 19-JUL-2000 WO 2000JP004862
 PR 19-JUL-1999 JP 99P 235910.24-MAR-2000 JP 00P 085377 PI
 AKIYO YAMADA, YOSHIIRO OZeki, TAKEO SAITO
 PC C1201/58.C12N15/29.C07K14/415, C07K16/16, C12P21/02.A01H5/00 CC
 FH Key Location/Qualifiers
 FT CDS (1). (1419).
 FEATURES
 source Location/Qualifiers
 1. .1602
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 1602; DB 6; Length 1602;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CACACCGTTCATTAACCAATGAAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60
 DB 1 CACACCGTTCATTAACCAATGAAGCTATGATGCTCGATTTCTCAAGCTTCTGATCTTGAC 60
 QY 61 AAGAAGAAAGCTCCTGAGATTTCTTCAATGCTTCGGCTCTTGAAGAAAGTCCCTTGG 120
 DB 61 AAGAAGAAAGCTCCTGAGATTTCTTCAATGCTTCGGCTCTTGAAGAAAGTCCCTTGG 120
 QY 121 GAACTTGGGGCTGGATTAATGCTGTTTACTGGTGAATGCTGAGAAAGCTGCCAGGTT 180
 DB 121 GAACTTGGGGCTGGATTAATGCTGTTTACTGGTGAATGCTGAGAAAGCTGCCAGGTT 180
 QY 181 ATTGCTCTGGAATTCATTTGAGAGTGTCTATCAAGAAAGTAAAGTAAATCAATGGGCACTAC 240
 DB 181 ATTGCTCTGGAATTCATTTGAGAGTGTCTATCAAGAAAGTAAAGTAAATCAATGGGCACTAC 240
 QY 241 AAAAATGTCAGCTTATGCTGCTGATGCTGATTTCTCCCACTCTCAGATTTCCACCAAT 300
 DB 241 AAAAATGTCAGCTTATGCTGCTGATGCTGATTTCTCCCACTCTCAGATTTCCACCAAT 300
 QY 301 TCAATGGATGATATTTCCCAATTTGTTACTCATGATATCTTCTGATGAAGAGGTGGAA 360
 DB 301 TCAATGGATGATATTTCCCAATTTGTTACTCATGATATCTTCTGATGAAGAGGTGGAA 360
 QY 361 AATTTGGTTGAAGAAATGTTGAATGTTGAAGCCAGGGGTTTACATTTTCTTCAGAGAA 420
 DB 361 AATTTGGTTGAAGAAATGTTGAATGTTGAAGCCAGGGGTTTACATTTTCTTCAGAGAA 420
 QY 421 TCTTGTTCATCAATCTGGGGATCAAAAGCAAAAGCAATCCCACTACCGTGAA 480
 DB 421 TCTTGTTCATCAATCTGGGGATCAAAAGCAAAAGCAATCCCACTACCGTGAA 480
 QY 481 CTTAGGTTCTACATPAAGGCTTCAAGAGTGTCTATTTGCAAGATGATCTGGAATCT 540
 DB 481 CTTAGGTTCTACATPAAGGCTTCAAGAGTGTCTATTTGCAAGATGATCTGGAATCT 540
 QY 541 TATGAGCTCTCCCTACTTAGCTGCAATGATTTGAGCTTATGCTGAGAACTCT 600
 DB 541 TATGAGCTCTCCCTACTTAGCTGCAATGATTTGAGCTTATGCTGAGAACTCT 600
 QY 601 CAGAACCAAGATAGTTGGTTGTGGCAAAAAGTTGATTTAAAGGATGATAAGGGGTTCCAG 660
 DB 601 CAGAACCAAGATAGTTGGTTGTGGCAAAAAGTTGATTTAAAGGATGATAAGGGGTTCCAG 660
 QY 661 CGATTTCTGGATCTAGCCAGTACAGTGAATAGCAATCTCGATATGAGCTGATTT 720
 DB 661 CGATTTCTGGATCTAGCCAGTACAGTGAATAGCAATCTCGATATGAGCTGATTT 720
 QY 721 GGCCCTGGTTATGTTAGCACTGGAGGATFAGAAACGACCAAGAGTGTGTCAATGCTG 780
 DB 721 GGCCCTGGTTATGTTAGCACTGGAGGATFAGAAACGACCAAGAGTGTGTCAATGCTG 780

QY 781 GACTTCGAAGCTGGCCAGAAAGTCTCGATGTTGGTGTGGAAATGGTGAGGTGACTTT 840
 DB 781 GACTTCGAAGCTGGCCAGAAAGTCTCGATGTTGGTGTGGAAATGGTGAGGTGACTTT 840
 QY 841 TACATGGCCGAGACCTTTGATGTTGAGGTGTTGAGATTTGATCTCTCCGTTAAATATGAT 900
 DB 841 TACATGGCCGAGACCTTTGATGTTGAGGTGTTGAGATTTGATCTCTCCGTTAAATATGAT 900
 QY 901 TCTTTTGGCCCTTGAAGCGTTCTATTGGGCTTAAATGTCCTGTTGAGTTTGAAGTAGCAGAT 960
 DB 901 TCTTTTGGCCCTTGAAGCGTTCTATTGGGCTTAAATGTCCTGTTGAGTTTGAAGTAGCAGAT 960
 QY 961 TGCACCAAGATAAATCACTCCCTGATACTCTTTGATGTCATCTATAGCCGTGACACCAT 1020
 DB 961 TGCACCAAGATAAATCACTCCCTGATACTCTTTGATGTCATCTATAGCCGTGACACCAT 1020
 QY 1021 CTGCATATTCAGGACAAGCTGCGTGTGTTAGATCTCTTACAAATGGTTGAAGCCAGCA 1080
 DB 1021 CTGCATATTCAGGACAAGCTGCGTGTGTTAGATCTCTTACAAATGGTTGAAGCCAGCA 1080
 QY 1081 GGTAAAGTTCTTAATCAGTGTATTCTGCAAGAAAGCTGCTCCAGCTCACTGAAATTCGCC 1140
 DB 1081 GGTAAAGTTCTTAATCAGTGTATTCTGCAAGAAAGCTGCTCCAGCTCACTGAAATTCGCC 1140
 QY 1141 GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGGAATATGCGCAGATGCTT 1200
 DB 1141 GCTTACATTAAGCAGAGGGGATATGATCTCCATGATGTAAGGAATATGCGCAGATGCTT 1200
 QY 1201 AAAGATGCTGGAATTTGTTGATGTTCTTGGCCGAGGATAGATCTCCATGATGTAAGGAATATGCGCAGAT 1260
 DB 1201 AAAGATGCTGGAATTTGTTGATGTTCTTGGCCGAGGATAGATCTCCATGATGTAAGGAATATGCGCAGAT 1260
 QY 1261 CTACCGAAGGAACTAGAGACTGTTGAGAGAAAGATGTTGTTCAATAGTATTTCTCT 1320
 DB 1261 CTACCGAAGGAACTAGAGACTGTTGAGAGAAAGATGTTGTTCAATAGTATTTCTCT 1320
 QY 1321 GAGGAGGATTAACATGACATTTGTTGAGGTTGGAATGATAAGTTGCGAGGACTGCCAAG 1380
 DB 1321 GAGGAGGATTAACATGACATTTGTTGAGGTTGGAATGATAAGTTGCGAGGACTGCCAAG 1380
 QY 1381 GGTGAGCAACGATGGGCTGCTGTTGTTCCCAAGAGAGTGAAGATCAGTTGCCGACT 1440
 DB 1381 GGTGAGCAACGATGGGCTGCTGTTGTTCCCAAGAGAGTGAAGATCAGTTGCCGACT 1440
 QY 1441 GGCATCTGCGATTTCTTGAATTAATCTTCAATGTTTTCATGTAATGTTCTTACATGT 1500
 DB 1441 GGCATCTGCGATTTCTTGAATTAATCTTCAATGTTTTCATGTAATGTTCTTACATGT 1500
 QY 1501 AAAATTTGCCAATAAGTTGCTATTTCCGAGACTGTAAGATGATTAATCATATTTTATCTTT 1560
 DB 1501 AAAATTTGCCAATAAGTTGCTATTTCCGAGACTGTAAGATGATTAATCATATTTTATCTTT 1560
 QY 1561 AATTAATCATGATTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602
 DB 1561 AATTAATCATGATTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602

RESULT 3
 AB080186
 LOCUS
 DEFINITION
 Sueoda japonica PEAMT mRNA for phosphoethanolamine
 N-methyltransferase, complete cds.
 ACCESSION
 AB080186
 VERSION
 AB080186.1 GI:28436073
 SOURCE
 Sueoda japonica
 ORGANISM
 Sueoda japonica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 Caryophyllales; Amaranthaceae; Sueoda.
 REFERENCE
 1 Yamada, A., Nozawa, G.T., Tanimoto, S. and Ozeki, Y.
 Glycinebetaine synthesis in Sueoda japonica

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1779)
AUTHORS Yamada, A., Norawa, G. T. and Ozeki, Y.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2002) Akiyo Yamada, Tokyo University of
Agriculture and Technology, Department of Biotechnology; Naka-cho
2-24-16, Koganei, Tokyo 184-8588, Japan
(E-mail: yamadenec@at.tuat.ac.jp, Tel: 81-423-88-7239,
Fax: 81-423-88-7239)
FEATURES Location/Qualifiers
source
1..1779
/organism="Suaeda japonica"
/mol_type="mRNA"
/db_xref="taxon:90346"
/country="Japan: Saga, Ariake Sea"
1..1779
/gene="PEAMT"
115..1599
/gene="PEAMT"
/codon_start=1
/product="phosphoethanolamine N-methyltransferase"
/protein_id="BAC57432.1"
/db_xref="GI:28436074"
/translation="MAASGMLQERDVFKKYVHEHTVLTIRAMLLDSQASDLQKE
RPEILSMPLPGLGKLLRGLAGIGRFTGELAEKAGOVIALDFIESAKNEVINGHYK
NYKFCADVTSPSPHSDLVFENMLWLSDEVNLSERMLKLPKGYIFER
ESCFQSGDHKKSNPHYRPFYTKAPEKCHLODGSNSVELSLSCCKIGATVRN
KNQOISLWOKVDSKODKGFQRLDTSQTKNSILRYRVEFGPVSYGTGHTKE
FVSMIDLKPGQKLVDSGIGGSGDFYMAETFDVEVGVFSDVSNMISFALERSIGLKA
VFPEVADCTKINYPDPSFDVTSRDITLHIDKPAFRSFKLPGKVLISDYCKK
AGPPSPFAAYIKQRYDLHDVKEYGQMLKADGPDVLAEDRTQPIRVLRKELETV
KEKDFVIFSDSEEDYNDIVGGNNDKLRRTAKGEQWGLFVAKKK"

ORIGIN
Query Match 100.0%; Score 1602; DB 8; Length 1779;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACACCGTTGATTACCAATGAAGCTGATGCTGATGCTCAAGCTTCGATCTTGAC 60
DB 178 CACACCGTTGATTACCAATGAAGCTGATGCTGATGCTCAAGCTTCGATCTTGAC 237
QY 61 AAGAGAGAGCTCTCAGATCTCTCAATGCTTCGGCTCTTGAAGAGAAATGCTCTTG 120
DB 238 AAGAGAGAGCTCTCAGATCTCTCAATGCTTCGGCTCTTGAAGAGAAATGCTCTTG 297
QY 121 GAATCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
DB 298 GAATCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
QY 181 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAAGTAATCAATGGGCACTAC 240
DB 358 ATTGCTCTGATTTCAATGAGAGTCTATCAAGAGAGATGAAGTAATCAATGGGCACTAC 417
QY 241 AAAAATGCTCAAGTTTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 418 AAAAATGCTCAAGTTTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
QY 301 TCATGATGATGATTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 478 TCATGATGATGATTTCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
QY 361 AATTGCTGTAAGAGATGTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 538 AATTGCTGTAAGAGATGTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
QY 421 TCTGTTTCCATCAATCGGGATCACAAACCAAGCAATCCCAACCACTACCGTGAA 480
DB 598 TCTGTTTCCATCAATCGGGATCACAAACCAAGCAATCCCAACCACTACCGTGAA 657
QY 481 CTTAGGTTCTACATAGGCTTCAAGAGTGTCTTTCGAGATGGAATCTGAAACTCT 540
DB 658 CTTAGGTTCTACATAGGCTTCAAGAGTGTCTTTCGAGATGGAATCTGAAACTCT 717

QY 541 TATGAGCTCTCCCTACTAGCTGCAATGTTATTTGGAGCTTATGTGAGAAACCAAGAAAAAC 600
DB 718 TATGAGCTCTCCCTACTAGCTGCAATGTTATTTGGAGCTTATGTGAGAAACCAAGAAAAAC 777
QY 601 CAGAACCAAGATTAGTTGGTGTGCAAAAGATTGATTCTTAAGGATGATAAGGGTTCCAG 660
DB 778 CAGAACCAAGATTAGTTGGTGTGCAAAAGATTGATTCTTAAGGATGATAAGGGTTCCAG 837
QY 661 CGATTTCTGATATCTAGCCAGTACAGCTGTAATAGCATTTCTGCGATGATGAGGCTGATTT 720
DB 838 CGATTTCTGATATCTAGCCAGTACAGCTGTAATAGCATTTCTGCGATGATGAGGCTGATTT 897
QY 721 GGCCTCTGGTTATGTTAGCACCTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG 780
DB 898 GGCCTCTGGTTATGTTAGCACCTGGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG 957
QY 781 GACTTGAAGCTGCGCAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
DB 958 GACTTGAAGCTGCGCAGAGGCTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1017
QY 841 TACATGCGGAGACCTTTTGTGATGTTGAGGTTGTTGATTTGATCTCTCCGTTAATATGAT 900
DB 1018 TACATGCGGAGACCTTTTGTGATGTTGAGGTTGTTGATTTGATCTCTCCGTTAATATGAT 1077
QY 901 TCCCTTTGCCCTTGGAGCTTCTAATTTGGGCTTAAATGTTCTGTTGAGTTTGGAGTAGCAGAT 960
DB 1078 TCCCTTTGCCCTTGGAGCTTCTAATTTGGGCTTAAATGTTCTGTTGAGTTTGGAGTAGCAGAT 1137
QY 961 TGACCAAGATTAACCTGATTAATCTTTTGAATCTGATCTATAGCGCTGACACCAT 1020
DB 1138 TGACCAAGATTAACCTGATTAATCTTTTGAATCTGATCTATAGCGCTGACACCAT 1197
QY 1021 CTGCATATTTCAAGCAAGCTGCTGTTGTTGATGCTCTTCTACAAATGTTGAAAGCCAGGA 1080
DB 1198 CTGCATATTTCAAGCAAGCTGCTGTTGTTGATGCTCTTCTACAAATGTTGAAAGCCAGGA 1257
QY 1081 GGTAAAGTTCTAATCAGTGATTTACTGCAAGAAAGCTGCTGCAACCTCACTGAAATCGCC 1140
DB 1258 GGTAAAGTTCTAATCAGTGATTTACTGCAAGAAAGCTGCTGCAACCTCACTGAAATCGCC 1317
QY 1141 GCTTACATTAAGCAGAGGGGATATGATCTCATGATGTAAGGAATATGGGAGTAGCTTT 1200
DB 1318 GCTTACATTAAGCAGAGGGGATATGATCTCATGATGTAAGGAATATGGGAGTAGCTTT 1377
QY 1201 AAAGATGCTGGAATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTTCATTTCGAGTT 1260
DB 1378 AAAGATGCTGGAATTTGTTGATGTTCTTCCGAGGATAGAACTGAGCAGTTTCATTTCGAGTT 1437
QY 1261 CTACCGAGGAGTACAGAGCTGTTGAGAGGAGGAGGATGTTTCATTAGTATTTCTCT 1320
DB 1438 CTACCGAGGAGTACAGAGCTGTTGAGAGGAGGAGGATGTTTCATTAGTATTTCTCT 1497
QY 1321 GAGAGGATTPACATGACATTTGTTGAGGTTGGAATGATTAAGTTGCGGAGGATGCGCAAG 1380
DB 1498 GAGAGGATTPACATGACATTTGTTGAGGTTGGAATGATTAAGTTGCGGAGGATGCGCAAG 1557
QY 1381 GGTGAGCAAGATGCGGCTCTGTTGTTGCCAAGAGAGATGAAGAACTCAGTTGCGGCACT 1440
DB 1558 GGTGAGCAAGATGCGGCTCTGTTGTTGCCAAGAGAGATGAAGAACTCAGTTGCGGCACT 1617
QY 1441 GGCACTGCTCAATTTCTAGTATTAATCTCAATGTTTTCATGTAATGATGTTCTACATGT 1500
DB 1618 GGCACTGCTCAATTTCTAGTATTAATCTCAATGTTTTCATGTAATGATGTTCTACATGT 1677
QY 1501 AAAATTGCCAATAGTTGCAATTTGCGAGATCTGTAAAGATGATTAATCATATTTATCTTTT 1560
DB 1678 AAAATTGCCAATAGTTGCAATTTGCGAGATCTGTAAAGATGATTAATCATATTTATCTTTT 1737
QY 1561 AATTAACTGAGTTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1602
DB 1738 AATTAACTGAGTTTATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAAATGCAAAAAA 1779


```

RESULT 4
AY133811
LOCUS
DEFINITION
Arabidopsis thaliana clone U1999 putative phosphoethanolamine
N-methyltransferase (At1g48600) mRNA, complete cds.
ACCESSION
AY133811
VERSION
AY133811.1 GI:22136801
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1459)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W.,
Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimn,P.,
Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1459)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W.,
Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K.,
Wong,C., Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H.,
Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A.,
Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M.,
Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Direct Submission
Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
Arabidopsis Full-length CDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K.,
Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Huan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,B.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimn,P.,
Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
Location/Qualifiers
1..1459
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="U1999"
/ecotype="Columbia"
note="This clone is in PUNI 51."
1..1459
/gene="At1g48600"
1..1428
/gene="At1g48600"
/evidence=experimental
/product="putative phosphoethanolamine"

```

```

N-methyltransferase"
/protein_id="AA091745.1"
/db_xref="GI:22136802"
/translation="MEHSDUTVAMMLDSKASDLKEPERVLSLIPYRGKSVLEL
GAGIQRFTGELAQAGEVIALDJFIESAIQKNSVNGHYKNIKFMCADVTSPDLKIKDG
SIDLIFSNLLMYLSDKEVELMAERMTGWKPGGYTFFRSCFHQSGDKRSKNPHTY
REPRFTYKVFQECQTRDASGNPSLSNGVCKICGAYVKNQKQNCQIOWKQVSVEND
KQDFRLDNVOYKSGSGLLYRVERFEGYSTGTFETKEFVAKMDLKPQKQVLDVDCG
IGGDFYMAENFDFHVVVICIDLSVNNMISFALERAIGLKCSVEFEVADCTTKTYPDNSFD
VYSRDTLTHIQDKPALPFTFKWLPKPGKVLITDYCRSAETSPSEAEYIKQGVYDL
HDVQYQGMKMDAGDDPVIADRTDQVQVLRLELKVKEKEKEEPISEDSSEYINDIV
GWSAKLRTASGSKWGLFIADKK"
misc_difference 282
/gene="At1g48600"
/note="compared to genomic sequence"
/replace="t"
1429..1459
/gene="At1g48600"

3'UTR
ORIGIN
Query Match 2.2% Score 35; DB 8; Length 1459;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 44 AGCTTCTGATTTGACAAAGACGCTCTGAG 78
Db 50 AAGCTTCTGATTTGACAAAGACGCTCTGAG 84

RESULT 5
AF428454
LOCUS
DEFINITION
Arabidopsis thaliana At1g48600/TIN15_20 mRNA, complete cds.
ACCESSION
AF428454
VERSION
AF428454.1 GI:16226648
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1656)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1656)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
Arabidopsis Full-length CDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the

```

sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H., Kim, C.J., Koesema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source
1. .1656
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL09-71-K18(R14923)"
/ecotype="Columbia"
1. .1446
/notes="putative phosphoethanolamine N-methyltransferase"
/codon_start=1
/product="At1g48600/TN15_20"
/protein_id="AAL16223.1"
/db_xref="GI:16226649"
/translation="MEHSSDLTVEAMWLDKASDLKKEERPEVLSPYPYEGKSVLEL
GAGIGRFTGELAQAGEVIALDFIESALQKNSVNGHYKNIKPCADVTSPDLKIKDG
SIDLIFSNWLLMYLSDKEVLAEMRMIGWPGGYIFFRESCHQSGDSKRNKPNTHY
KDFQFPLDQVQSCQTRDASGNSFELSMVCKICIGAYVKNKQNCQICWIKQVSVEND
IGGGDFLDMVQVSGIILRYVFGVSTGPGFTTKEFVAKMDLPGQKQVLDVCGG
VYVSRDILHIQKALPRTFFKWLKPGKVLITDYCRSAETSPSPFAEYIKQGYDL
HDVQAYGQMLKADGDDVIAEDRTDQFQVLRLELKEKEKEBEFISDFSEEDYNDIV
GWSAKLERTASGEQKWLFTADKK"
1447. .1656

5'UTR

CDS

ORIGIN

Query Match 2.2%; Score 35; DB 8; Length 1656;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 78
|||||
DB 68 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 102
|||||

RESULT 6

AY063866

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ORIGIN

FEATURES

source

gene

5'UTR

CDS

misc_difference

3'UTR

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Score

DB

Length

Pred. No.

3.7e-07

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (19-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Location/Qualifiers
1. .1660
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL06-86-D12 (R11999)"
/ecotype="Columbia"
/note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SstI insert."
1. .1660
/genes="At1g48600"
1. .34
/genes="At1g48600"
35. .1462
/genes="At1g48600"
/codon_start=1
/evidence="experimental"
/product="putative phosphoethanolamine N-methyltransferase"
/protein_id="AAL136222.1"
/db_xref="GI:17380784"
/translation="MEHSSDLTVEAMWLDKASDLKKEERPEVLSPYPYEGKSVLEL
GAGIGRFTGELAQAGEVIALDFIESALQKNSVNGHYKNIKPCADVTSPDLKIKDG
SIDLIFSNWLLMYLSDKEVLAEMRMIGWPGGYIFFRESCHQSGDSKRNKPNTHY
KDFQFPLDQVQSCQTRDASGNSFELSMVCKICIGAYVKNKQNCQICWIKQVSVEND
IGGGDFLDMVQVSGIILRYVFGVSTGPGFTTKEFVAKMDLPGQKQVLDVCGG
VYVSRDILHIQKALPRTFFKWLKPGKVLITDYCRSAETSPSPFAEYIKQGYDL
HDVQAYGQMLKADGDDVIAEDRTDQFQVLRLELKEKEKEBEFISDFSEEDYNDIV
GWSAKLERTASGEQKWLFTADKK"

misc_difference 316
/genes="At1g48600"
/notes="compared to genomic sequence"
/replaces="t"
1463. .1660
/genes="At1g48600"

3'UTR
ORIGIN

Query Match 2.2%; Score 35; DB 8; Length 1660;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS
F*
join(21967..22130,22278..22378)
/note="unknown protein; similar to ESTs gb|AF162845.1,
gb|AI992553.1, and gb|AA042117.1"
/codon start=1
/evidence=not experimental
/product="TIN15.5"
/protein_id="AAF79693.1"
/db_xref="GI:8778685"
/translation="MALQWLILSYVAVEVITLVLPMLKKRVQLVSLILQ
ASIVAFQQLDIYKAEHLSCSSCEVATERDYKSVSTIP"
complement(join(23638..2481,24564..24859,25106..25237))
/note="unknown protein; similar to EST gb|AI996045.1"
/codon start=1
/evidence=not experimental
/product="TIN15.6"
/protein_id="AAF79715.1"
/db_xref="GI:8778707"
/translation="MDAVAMLPQSPILRLCTSERLJESTPCFGRMISGGPQPELR
QNSLKPFPQGTCTCVKRSFVYKASGADSTESIAFLQLKSVKQFLQILVS
HPLVPAVEQQLQIDRDABEQSKASSVLGTDIVLIRIAEYKERRALEEI
LYALVQKMDANVTLPVPSITSSADPSGRVDTPMTLDGELRLHSPVTEMIONHLS
IILKNTDITVAQISKLGVQVYASVYGYFLKIDQRFQLEKTRILPGSGDEG
ETSEIQAGRDVERNFYEABETTYQAVSNQDVSGVGINASGGSFDMKQSKLTYV
MSPGGLTQRYATIRSESGVIEKHEALFCEPEIVITPQGTITDSSKDHKIKFKG
LKLVLAVTPGSLFWDVSHVSRHPVLN"
join(26361..26450,26546..26588,26676..27014,27366..27706,
27823..28041)
/note="unknown protein; similar to ESTs emb|Z33851.1,
gb|Z33009.1, and gb|AA94999.1"
/codon start=1
/evidence=not experimental
/product="TIN15.7"
/protein_id="AAF79694.1"
/db_xref="GI:8778686"
/translation="MESKALICLFLPRLRSSPRLLSPSPASSTFATRHKLDSR
QTLMLKPOLSRVRSQSSQSDRPEKQKQSNYARAEFLPQKSGSVGFLGTLQ
VERSLKAPROEKGFTVLAFLVLLSILPQFLSGIIEATFNQDVAGRSIV
TSCEFTVYAGLAIELSVTDVQRPYLDPSKMGILITGLRGVLTAPITMGLKVVY
PFAVNTFALGIDALIANFLPVGCAVQRPPEARLERSGSCWPIVIFVEVRLY
QVTRATFVQLMFMMDAATTAETIRGVALVGLVTLQFLAVNCLWSFITFLMLF
PSRVEGNY"
join(28676..28749,28903..28942,29019..29111,29232..29273,
29360..29466,29687..29774,29870..30073,30240..30293,
30397..30594,30675..30818)
/note="similar to glutamine synthetase cytosolic isozyme 2
sp|P5119|GLN2 VITVI; similar to EST emb|Z29893.1"
/codon start=1
/evidence=not experimental
/product="TIN15.8"
/protein_id="AAF79695.1"
/db_xref="GI:8778687"
/translation="MTSPSLDLNLDLSTKIIARIYIWIGSGMDIRSKARTLPGPV
SNPTKLPKNYDSSDQAGDDSEAFKDPFRKNNILVCMDAYRAGDDIPFNH
KAVIFDHPNABEPMFGIEQYTLKKDVKWPLGFPQPGQFYCAVADKA
PGRIDVHAKYACLSGLSIGGAEVPGQWFOISPTVIGIGAGDQLVARYILERI
TEICGVTVSPDKPIQGDWNGAAAHFTPKMKRKGGLDLIKAKIKLEVKKHQIA
AYGGERRLTGKHETADINTPMSGVADRGASVRVGRDTEKEGKGYFEDRPPSNMOP
YLVTSMIAETIL"
join(33367..34735,35037..35635)
/note="similar to receptor-like protein kinase
2.23; Score 35; DB 8; Length 95295;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 44 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 78
DB 81389 AAGCTTCTGATCTTGACAAAGAGACGCTCTGAG 81423

LOCUS AX250553 1120 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 3 from Patent WO0168870.
ACCESSION AX250553
VERSION AX250553.1 GI:15984290
KEYWORDS
SOURCE Spinacia oleracea (spinach)
ORGANISM Spinacia oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Spinacia.
1
REFERENCE
AUTHORS Hanson, A.D., Nuccio, M.L. and Henry, S.A.
TITLE S-adenosyl-L-methionine:phosphoethanolamine n-methyltransferase
compositions and methods for modulating lipid biosynthesis in
plants
JOURNAL Patent: WO 0168870-A 3 20-SEP-2001;
University of Florida (US); Carnegie-Mellon University (US)
FEATURES
source
1..1120
/organism="Spinacia oleracea"
/mol_type="unassigned DNA"
/db_xref="taxon:3562"
ORIGIN
Query Match 2.0%; Score 32; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 GCTGGTATTGCTGCTTTTACTGCTGAATGGC 161
DB 446 GCTGGTATTGCTGCTTTTACTGCTGAATGGC 477
RESULT 9
LOCUS AX250551 2235 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 1 from Patent WO0168870.
ACCESSION AX250551
VERSION AX250551.1 GI:15984289
KEYWORDS
SOURCE Spinacia oleracea (spinach)
ORGANISM Spinacia oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Spinacia.
1
REFERENCE
AUTHORS Hanson, A.D., Nuccio, M.L. and Henry, S.A.
TITLE S-adenosyl-L-methionine:phosphoethanolamine n-methyltransferase
compositions and methods for modulating lipid biosynthesis in
plants
JOURNAL Patent: WO 0168870-A 1 20-SEP-2001;
University of Florida (US); Carnegie-Mellon University (US)
FEATURES
source
1..2235
/organism="Spinacia oleracea"
/mol_type="unassigned DNA"
/db_xref="taxon:3562"
ORIGIN
Query Match 2.0%; Score 32; DB 6; Length 2235;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 GCTGGTATTGCTGCTTTTACTGCTGAATGGC 161
DB 446 GCTGGTATTGCTGCTTTTACTGCTGAATGGC 477
RESULT 10
LOCUS AF237633 2552 bp mRNA linear PCN 10-MAY-2000
DEFINITION Spinacia oleracea phosphoethanolamine N-methyltransferase (PEAMT)
mRNA, complete cds.

ACCESSION AF2337633
 VERSION AF2337633.1 GI:7407188
 KEYWORDS
 SOURCE Spinacia oleracea (spinach)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Anaranthaceae; Spinacia.
 REFERENCE 1 (bases 1 to 2252)
 AUTHORS Nuccio, M.L., Ziemak, M.J., Henry, S.A., Weretilnyk, E.A. and Hanson, A.D.
 TITLE cDNA cloning of phosphoethanolamine N-methyltransferase from spinach by complementation in *Schizosaccharomyces pombe* and characterization of the recombinant enzyme
 JOURNAL J. Biol. Chem. 275 (19), 14095-14101 (2000)
 MEDLINE 20261526
 PUBMED 10799484
 REFERENCE 2 (bases 1 to 2252)
 AUTHORS Nuccio, M.L. and Hanson, A.D.
 TITLE Direct Submission
 JOURNAL Submitted (21-PEB-2000) Horticultural Sciences, University of Florida, P.O. Box 110690, Gainesville, FL 32611, USA
 FEATURES
 source
 1..2252
 /organism="Spinacia oleracea"
 /mol_type="mRNA"
 /db_xref="taxon:3562"
 /cissue_type="salinized leaf"
 1..2252
 /gene="PEAMT"
 284..1738
 /gene="PEAMT"
 /EC_number="2.1.1.103"
 /function="catalyzes S-adenosylmethionine dependent N-methylation of phosphoethanolamine, phosphonomethylmethanolamine and phosphodimethylmethanolamine"
 /codon_start=1
 /product="phosphoethanolamine N-methyltransferase"
 /protein_id="AAF61950.1"
 /db_xref="GI:7407189"
 /translation="MAAGACVQLQREVFKKYYWHSVLTVEANMLDSQASDLKVE RPEVLSMLPPYEGKSVLGLAGIGRPTGLAEKASQVIALDFIESVKNESINGHYK NVKFCADVTSPISNFSNVDIIFSNLLMYLSDEVERLVMKMLKPGGYIFPR ESCFQSGDHRKKNPHYRPFRTKIFKECHQMDSGNSYELSLGCKIGAVVKS KKNQKISLWQKVSDDKGFQFLDSQYKFNLSLYRVRVFGYVSTGLSTTKE FVSKLDLPGQKVLVDGCGIGGDFYMAENVYDVGVGIDLSINMISPALESGICLKA VEFVADCTKKQYDPSFNVYSRDTLHLDKPALEFRPHKWLAPGKVLISDYCKS ACTPSAEPAAVIRQGVLDHVKAGKMLKAGFVEVIAENRTDQFIQVLKELDALE QEKDFIDFSEEDYNDLVGNKALVRITTEGSEQQGLFIKKM"
 ORIGIN
 Query Match 2.0%; Score 32; DB 8; Length 2252;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 130 GCTGGTATGGTCGTTTACTGGTGGAATTGCG 161
 DB 446 GCTGGTATGGTCGTTTACTGGTGGAATTGCG 477
 RESULT 11
 BC049504
 LOCUS
 DEFINITION Danto rerio mRNA similar to estrogen receptor-binding fragment-associated gene 9 (cDNA clone MGC:56608 IMAGE:5915434), complete cds.
 ACCESSION BC049504
 VERSION BC049504.1 GI:29436967
 KEYWORDS MGC.
 SOURCE Danto rerio (zebrafish)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE AUTHORS

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 1347)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Skapleton, L., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwen, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vellala, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1347)
 Strausberg, R.
 Direct Submission
 Submitted (31-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Sumio Sugano
 cDNA Library Preparation: Dr. Sumio Sugano
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nese, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

TITLE

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL

REMARK COMMENT

FEATURES source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 104 Row: 1 Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
 Location/Qualifiers
 1..1347
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:56608 IMAGE:5915434"
 /tissue_type="Whole body, adult male"
 /clone_lib="Sugano SJD adult male"
 /lab_host="DH10B"
 /note="Vector: pME18S-FL3"
 207..845
 /codon_start=1
 /product="Similar to estrogen receptor-binding fragment-associated gene 9"
 /protein_id="AA049504.1"
 /db_xref="GI:29436968"

CDS

/translation="MAITQRLFKICTCLASILSPFKRLICRGRSLSGDITLPT
TVYSSASKPEIEMSSWDEAPTSIKIEGNGIVAPPQNOAEEBPDYKDMAPT
RTKIVLKKEPINFMSPPSSGFSRLAATODMSFIOPASLGLDPTWODNNAM
DEADAWEAEVLRQKAEEREMQKQKEVQRMKKQKIAVKLS"

ORIGIN

Query Match 1.9%; Score 31; DB 5; Length 1347;
Best Local Similarity 100.0%; Pred. No. 5.1e-05; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 0;

Oy 1572 GATTATGCAAAAAA...AAAAA 1602
Db 1295 GATTATGCAAAAAA...AAAAA 1325

RESULT 12
BX324004/c 138116 bp DNA linear HTG 24-SEP-2003
LOCUS
DEFINITION Danio rerio clone CH211-284017, WORKING DRAFT SEQUENCE, 4 unordered
pieces.

ACCESSION BX324004
VERSION BX324004.3 GI:35209309
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)

ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 138116)

REFERENCE
AUTHORS McLaren, S.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 24, 2003 this sequence version replaced gi:30349842.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC284017
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 137487 bases at least Q40
Consensus quality: 137601 bases at least Q30
Consensus quality: 137724 bases at least Q20
Insert size: 137816; sum-of-contigs
Quality coverage: 8.58x in Q20 bases; sum-of-contigs Quality
coverage: 8.29x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4524: contig of 4524 bp in length
* 4525 4624: gap of 100 bp
* 4625 39932: contig of 35308 bp in length
* 39933 40032: gap of 100 bp
* 40033 101266: contig of 61234 bp in length
* 101267 101366: gap of 100 bp
* 101367 138116: contig of 36750 bp in length.

Location/Qualifiers
1...138116
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"

FEATURES

source
1...138116
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"

/clone="CH211-284017"
/clone_lib="CHORI-211"
1...4524
/note="assembly fragment:02039
fragment chain:1"

misc_feature

4625...39932
/note="assembly fragment:01037
fragment chain:1
clone end:T7
vector_side:right"

misc_feature

40033...101266
/note="assembly fragment:00551
fragment chain:1
clone end:SP6
vector_side:left"

misc_feature

101367...138116
/note="assembly fragment:00693
fragment chain:1"

ORIGIN

Query Match 1.9%; Score 31; DB 2; Length 138116;
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 0;

Oy 1572 GATTATGCAAAAAA...AAAAA 1602
Db 62624 GATTATGCAAAAAA...AAAAA 62594

RESULT 13

LOCUS
DEFINITION BX293550 168179 bp DNA linear HTG 02-JUN-2003
Danio rerio clone CH211-276D14, WORKING DRAFT SEQUENCE, 2 unordered
pieces.

ACCESSION

BX293550
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 168179)

REFERENCE

AUTHORS McLaren, S.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 23, 2003 this sequence version replaced gi:29500500.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC276D14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 167777 bases at least Q40
Consensus quality: 167817 bases at least Q30
Consensus quality: 167839 bases at least Q20
Insert size: 168079; sum-of-contigs
Quality coverage: 9.06x in Q20 bases; sum-of-contigs Quality
coverage: 8.14x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 4524: contig of 4524 bp in length
* 4525 4624: gap of 100 bp
* 4625 39932: contig of 35308 bp in length
* 39933 40032: gap of 100 bp
* 40033 101266: contig of 61234 bp in length
* 101267 101366: gap of 100 bp
* 101367 138116: contig of 36750 bp in length.

* as soon as it is available and the accession number will
* be preserved.

1 152355: contig of 152355 bp in length
152356 152455: gap of 100 bp
152456 168179: contig of 15724 bp in length.

FEATURES

source

Location/Qualifiers
1..168179
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-276D14"
/clone_lib="CHORI-211"
1..152355
/notes="assembly fragment:00072
fragment_chain:1
clone_end:SP6
vector_side:left"
152455..168179
/notes="assembly_fragment:01003
fragment_chain:1
clone_end:T7
vector_side:right"

misc_feature

fragment_chain:1
clone_end:SP6

misc_feature

152455..168179
/notes="assembly_fragment:01003
fragment_chain:1
clone_end:T7
vector_side:right"

ORIGIN

Query Match 1.9%; Score 31; DB 2; Length 168179;
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
|||||
DB 45936 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 45966

RESULT 14

BX005125/c

LOCUS

DEFINITION Zebrafish DNA sequence from clone CH211-127116 in linkage group 10,
complete sequence.

BX005125

VERSION

BX005125.11 GI:37699970

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 172587)

Beasley H.

Direct Submission

TITLE

JOURNAL

Submitted (16-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Oct 16, 2003 this sequence version replaced gi:34221835.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compression-and-repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhirong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/projects/d_rerio/fishmask.shtm

CH211-127116 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES

source

Location/Qualifiers
1..172587
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-127116"
/clone_lib="CHORI-211"

ORIGIN

Query Match 1.9%; Score 31; DB 5; Length 172587;
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
|||||
DB 48349 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 48319

RESULT 15

BX530060

LOCUS

DEFINITION

Danio rerio clone DKEX-283P16, WORKING DRAFT SEQUENCE, 3 unordered
pieces.

BX530060

VERSION

BX530060.3 GI:35209357

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 175745)

McLaren, S.

Direct Submission

TITLE

JOURNAL

Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Sep 24, 2003 this sequence version replaced gi:31408112.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk

----- Project Information

Center project name: zK283P16

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 175264 bases at least Q40

Consensus quality: 175360 bases at least Q30

Consensus quality: 175437 bases at least Q20

Insert size: 175545; sum-of-contigs

Insert size: 18158; 1.6% error; agarose-fp

Quality coverage: 9.45x in Q20 bases; sum-of-contigs Quality

Coverage: 9.06x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

22689: contig of 22689 bp in length
22690: gap of 100 bp
22790: contig of 83479 bp in length
106268: contig of 106268 bp in length
106369: gap of 100 bp
175745: contig of 69377 bp in length.

FEATURES

source

1. 175745
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEX-283F16"
/clone_lib="DanioKey"

misc_feature

1. 22689
/note="assembly fragment:01399
fragment_chain:1"

misc_feature

22790..106268
/note="assembly fragment:02444
fragment_chain:1"

misc_feature

106369..175745
/note="assembly fragment:02005.0"

ORIGIN

Query Match 1.9%; Score 31; DB 2; Length 175745;
Best Local Similarity 100.0%; Pred.No. 3.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1572 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 110763 GATTATGCAAAAAAAAAAAAAAAAAAAAAA 110793

Search completed: August 2, 2004, 13:13:09
Job time : 6485 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 08:17:48 ; Search time 668 Seconds

(without alignments)
10188.044 Million cell updates/sec

Title: US-10-031-331B-39

Perfect score: 1602

Sequence: 1 cacacgctgattacacat.....aaaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO_NVC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	100.0	1602	4	Aaf74206 DNA encod
2	32	2.0	1120	5	Aad16798 Spinach p
3	32	2.0	2235	5	Aad16797 Spinach p
4	29	1.8	438	4	Aak56740 Human imm
5	29	1.8	483	9	Ade82102 Arabidops
6	29	1.8	507	7	Abz20190 Group III
7	29	1.8	536	4	Aak64289 Human imm
8	29	1.8	1267	4	Aas26138 Human cdn
9	29	1.8	1267	7	Abx73479 Human nov
10	28	1.7	307	6	Ab186191 Human ova
11	28	1.7	339	4	Aas60367 Human can
12	28	1.7	419	6	Abt09151 Phase-1 R
13	28	1.7	450	4	Abas7560 Human foe
14	28	1.7	450	4	Aai37117 Probe #58
15	28	1.7	450	4	Abaz7022 Probe #54
16	28	1.7	450	4	Aak31213 Human bon
17	28	1.7	450	4	Aak05609 Human bra
18	28	1.7	450	4	Abz30894 Human liv
19	28	1.7	450	6	Abz05965 Human gen
20	28	1.7	559	6	Abv78106 Hypoxia-r
21	28	1.7	1417	6	Abx04582 Human end
22	28	1.7	2000	6	Abz16984 Arabidops
23	28	1.7	3183	6	Abz64605 Human cdn

C 24 28 1-7 10221 4 AAF58417 Human oli
C 25 28 1-7 11655 4 AAK71335 Human imm
C 26 28 1-7 45121 8 ADA02744 Human TNF
C 27 28 1-7 45121 9 ADE72482 Human TNF
C 28 28 1-7 45121 9 ADC85224 Human TNF
C 29 28 1-7 42 7 ABT23108 Library c
C 30 27 1-7 42 7 ABT23109 Library c
C 31 27 1-7 42 7 ABT23110 Library c
C 32 27 1-7 43 7 ABT23111 Library c
C 33 27 1-7 182 2 AAO68945 Junction
C 34 27 1-7 260 9 ADD32852 Human mit
C 35 27 1-7 444 6 AB164210 Stomach c
C 36 27 1-7 444 6 AB167975 Ovary can
C 37 27 1-7 444 6 ABM96383 Gene #288
C 38 27 1-7 491 6 ABQ57289 Human col
C 39 27 1-7 513 8 ACH34265 Human end
C 40 27 1-7 752 5 AAS69783 DNA encod
C 41 27 1-7 804 4 AAL23782 Human bre
C 42 27 1-7 978 4 AAH31349 Human sec
C 43 27 1-7 978 7 ADA56332 Gene enco
C 44 27 1-7 978 7 ADA40173 Human sec
C 45 27 1-7 978 9 ADC73718 Human sec

ALIGNMENTS

RESULT 1
ID AAF74206
XX AAF74206 standard; DNA; 1602 BP.
AC AAF74206;
DT 06-AUG-2003 (revised)
DT 02-MAY-2001 (first entry)
DE DNA encoding environmental stress tolerant protein SEQ ID 39.
XX Environmental stress resistance; salt; heat; desert; transgenic plant;
KW ds.
XX Suaeda japonica.
XX WO200106006-A1.
XX 25-JAN-2001.
XX 19-JUL-2000; 2000WO-JP004862.
XX 19-JUL-1999; 99JP-00235910.
XX 24-MAR-2000; 2000JP-00085377.
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX Yamada A, Ozeki Y, Saito T;
XX WPI; 2001-147355/15.
XX P-PSDB; AAB80627.
XX Screening method to obtain DNA encoding environmental stress resistance factor, useful for producing transgenic plants resistant to environmental stress.
XX Claim 65; Page 123-125; 167pp; Japanese.
XX Polynucleotide sequences AAF74187 - AAF74218 encode proteins AAB80608 - AAB80639, which impart environmental stress resistance. The invention relates to a method for identifying DNA encoding proteins imparting environmental stress resistance. The method comprises inserting cDNA from a library originating in a salt-resistant organism into a host cell, culturing the transformants under conditions in which the untransformed host does not grow well, and selecting for viable clones. The method is useful for obtaining environmental stress resistance

CC factors. The DNA encoding proteins conferring environmental stress resistance can be used in the production of plants resistant to environmental stress, which can be cultivated in unfavourable environments such as deserts, salt damaged ground, cold regions and the oceans. They can be used for increasing the area of land covered by green plants, and desert greening and afforestation, in order to counter the effects of the increase in atmospheric carbon dioxide concentration. PCR primers AAF74219 and AAF74220 are used in an example illustrating the method of the invention. (Updated on 06-AUG-2003 to correct OS field.)		
CC	XX	Sequence 1602 BP; 454 A; 279 C; 383 G; 486 T; 0 U; 0 Other;
CC	XX	Query Match 100.0%; Score 1602; DB 4; Length 1602;
CC	XX	Best Local Similarity 100.0%; Pred. No. 0;
CC	XX	Matches 1602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	DB	1 CACACGGTGGATTAAACCAATGAAGCTATGATGCTCCGATCTCAAGCTTCTGATCTTGAC 60
DB	DB	1 CACACGGTGGATTAAACCAATGAAGCTATGATGCTCCGATCTCAAGCTTCTGATCTTGAC 60
QY	DB	61 AAAGAAAGACGCTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG 120
DB	DB	61 AAAGAAAGACGCTCTGAGATCTTTCAATGCTTCGGCTCTTGAAGGAAATGCTCTTG 120
QY	DB	121 GAATCTGGGCTGGTATGCTGTTTACTGTGATGCTGAAAGCTGCGCAGGTT 180
DB	DB	121 GAATCTGGGCTGGTATGCTGTTTACTGTGATGCTGAAAGCTGCGCAGGTT 180
QY	DB	181 ATTGCTCTGATTTCAATGAGTGTCTATCAAGAGATGAAGTAATCAATGGCCTAC 240
DB	DB	181 ATTGCTCTGATTTCAATGAGTGTCTATCAAGAGATGAAGTAATCAATGGCCTAC 240
QY	DB	241 AAAAAATGCAAGTTTATGCTGCTGATGTGATCTTCCCACTCTCAGTTTCCCAACCAT 300
DB	DB	241 AAAAAATGCAAGTTTATGCTGCTGATGTGATCTTCCCACTCTCAGTTTCCCAACCAT 300
QY	DB	301 TCATTGATGATATCTCCAAATGGTTTACTCATGTATCTTCTGATGAGAGGTGGA 360
DB	DB	301 TCATTGATGATATCTCCAAATGGTTTACTCATGTATCTTCTGATGAGAGGTGGA 360
QY	DB	361 AATTGGTTCGAAAGATGTTGAAAGCTTGAAGCAGGGGGTTACATTTCTTCAGAGAA 420
DB	DB	361 AATTGGTTCGAAAGATGTTGAAAGCTTGAAGCAGGGGGTTACATTTCTTCAGAGAA 420
QY	DB	421 TCTTTTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCAACCACTACCGTGAA 480
DB	DB	421 TCTTTTTCATCAATCTGGGATCAAAACGCAAAAGCAATCCCAACCACTACCGTGAA 480
QY	DB	481 CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCTTTGCAAGATGGATCTGGAACCTCT 540
DB	DB	481 CCTAGGTTCTACACTAAGGCTTCAAGAGTGTCTTTGCAAGATGGATCTGGAACCTCT 540
QY	DB	541 TATGAGCTCTCCCTACTAGCTGCAATGTTATGGAGCTTATGTTCAGAAACAGAAAAAC 600
DB	DB	541 TATGAGCTCTCCCTACTAGCTGCAATGTTATGGAGCTTATGTTCAGAAACAGAAAAAC 600
QY	DB	601 CAGAACAGATGATGCTGTTGTCGCAAAAGTTGATTTCTAAGGATGATAAGGGTTCAG 660
DB	DB	601 CAGAACAGATGATGCTGTTGTCGCAAAAGTTGATTTCTAAGGATGATAAGGGTTCAG 660
QY	DB	661 CGATTTCTGGTACTAGCCAGTACAAAGTGTAAATAGCAATCTTCGGATATGAGCGTATT 720
DB	DB	661 CGATTTCTGGTACTAGCCAGTACAAAGTGTAAATAGCAATCTTCGGATATGAGCGTATT 720
QY	DB	721 GGCCTGTTGATTTAGTACCTGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG 780
DB	DB	721 GGCCTGTTGATTTAGTACCTGAGGATATGAACCAACCAAGAGTTTGTGTCAATGCTG 780
QY	DB	781 GACTTGAAGCTGGCCAGAGGCTCTGATGTTGTTGGTGTGGAATGGTGGAGTCACTTT 840
DB	DB	781 GACTTGAAGCTGGCCAGAGGCTCTGATGTTGTTGGTGTGGAATGGTGGAGTCACTTT 840
QY	DB	841 TATACGCGGAGACCTTTGATGTTGAGGTTGTTGGATTTGATCTCTCCGTTAATATGATT 900

RESULT 2

AAD16798 standard; DNA; 1120 BP.

XX AAD16798;

XX AC

XX 29-NOV-2001 (first entry)

XX Spinach PEAMT truncated DNA.

XX Spinach; S-adenosyl-L-methionine; phosphoethanolamine N-methyltransferase; PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline;

XX phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine;

XX glycine betaine; choline-O-sulphate; lipid content alteration;

XX osmotic stress tolerance; nutritional value; transgenic plant;

XX cryoprotectant; ds.

XX Spinacia oleracea.

XX OS Location/Qualifiers

XX Key 254..1120

XX FH

XX CDS

FT /*tag= a
 FT /product= "Spinach PEAMT truncated protein"
 FT /note= "CDS does not include stop codon"
 FT /partial

WO200168870-A2.

20-SEP-2001.

15-MAR-2001; 2001WO-US008352.

15-MAR-2000; 2000US-00525885.

(UYFL) UNIV FLORIDA.

(UYCA-) UNIV CARNEGIE MELLON.

Hanson AD, Nuccio ML, Henry SA;

WPI; 2001-565796/63.

P-PSDB; AAE09761.

New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase polypeptides, useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine and for altering the lipid content in plants cells.

Claim 18; Page 109; 158pp; English.

The present sequence is spinach S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase (PEAMT) truncated DNA. The PEAMT sequences are useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine, phosphomono-methylethanolamine, choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine betaine. They are useful for altering the lipid content in plant cells. The polynucleotides are also useful for improving the osmotic stress tolerance of a plant and increasing the cryoprotectant properties of a plant. The present invention also relates to methods and compositions comprising PEAMT used for generating transgenic plants with increased nutritional value

Sequence 1120 BP; 304 A; 218 C; 256 G; 342 T; 0 U; 0 Other;

Query Match 2.0%; Score 32; DB 5; Length 1120;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GCTGGTATTGGTCGTTTACTGGTGAATTGCG 161

Db 446 GCTGGTATTGGTCGTTTACTGGTGAATTGCG 477

RESULT 3

AA016797

ID AAD16797 standard; DNA; 2235 BP.

XX AC

XX AAD16797;

DT 29-NOV-2001 (first entry)

DE Spinach PEAMT DNA.

XX XX

Spinach; S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase; PEAMT; cellular intermediate; phospho-dimethylethanolamine; choline; phosphomono-methylethanolamine; phosphocholine; phosphatidylcholine; glycine betaine; choline-O-sulphate; lipid content alteration; osmotic stress tolerance; nutritional value; transgenic plant; cryoprotectant; db.

OS Spinacia oleracea.

XX XX

FT Key Location/Qualifiers

FT CDS 254..1738

FT /*tag= a

FT /product= "Spinach PEAMT protein"

WO200168870-A2.

20-SEP-2001.

15-MAR-2001; 2001WO-US008352.

15-MAR-2000; 2000US-00525885.

(UYFL) UNIV FLORIDA.

(UYCA-) UNIV CARNEGIE MELLON.

Hanson AD, Nuccio ML, Henry SA;

WPI; 2001-565796/63.

P-PSDB; AAE09760.

New plant S-adenosyl-L-methionine:phosphoethanolamine N-methyltransferase polypeptides, useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine and for altering the lipid content in plants cells.

Claim 18; Page 108-109; 158pp; English.

The present sequence is a DNA encoding spinach S-adenosyl-L-methionine: phosphoethanolamine N-methyltransferase (PEAMT). The PEAMT sequences are useful for modulating the levels of cellular intermediates such as phosphodimethylethanolamine, phosphomono-methylethanolamine, choline, phosphocholine, phosphatidylcholine, choline-O-sulphate or glycine betaine. They are useful for altering the lipid content in plant cells. The polynucleotides are also useful for improving the osmotic stress tolerance of a plant and increasing the cryoprotectant properties of a plant. The present invention also relates to methods and compositions comprising PEAMT used for generating transgenic plants with increased nutritional value

Sequence 2235 BP; 623 A; 399 C; 516 G; 697 T; 0 U; 0 Other;

Query Match 2.0%; Score 32; DB 5; Length 2235;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GCTGGTATTGGTCGTTTACTGGTGAATTGCG 161

Db 446 GCTGGTATTGGTCGTTTACTGGTGAATTGCG 477

RESULT 4.

AAK56740

ID AAK56740 standard; cDNA; 438 BP.

XX AC

XX AAK56740;

DT 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1800.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ss.

OS Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179085P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205513P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0211886P.
PR 30-JUN-2000; 2000US-0215133P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229609P.
PR 06-SEP-2000; 2000US-0229913P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231342P.
PR 08-SEP-2000; 2000US-0231343P.
PR 08-SEP-2000; 2000US-0231344P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0233223P.
PR 21-SEP-2000; 2000US-0233224P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-483426/52.
P-PSDB; AAM83959.
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
useful for preventing, diagnosing and/or treating cancers and metastasis.
Claim 1; SEQ ID NO 1800; 3071pp + Sequence Listing; English.
AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)

CC amino acid sequences given in AAK821170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK92169
 CC represent sequences used in the exemplification of the present invention
 XX

SQ Sequence 438 BP; 177 A; 63 C; 67 G; 128 T; 0 U; 3 Other;
 Query Match 1.8%; Score 29; DB 4; Length 438;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
 Db 397 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 425

RESULT 5

ID ADE82102 standard; cDNA; 483 BP.

XX AC ADE82102;

XX DT 29-JAN-2004 (first entry)

XX DE Arabidopsis thaliana expressed polynucleotide seq id 873.

XX KW Genetically modified organism; transgenic organism; plant;
 KW inhibitor testing; activator testing; modifier testing; fungicide;
 KW insecticide; genetic function; genetic regulation; cellular metabolism;
 KW gene; ss.

XX OS Arabidopsis thaliana.

XX FN US2003115639-A1.

XX PD 19-JUN-2003.

XX PF 26-JAN-2001; 2001US-00770961.

XX PR 27-JAN-2000; 2000US-0178466P.

XX PA (GORL/) GORLACH J.
 XX PA (ANY/) AN Y.
 XX PA (HAMI/) HAMILTON C M.
 XX PA (PRIC/) PRICE J L.
 XX PA (RAIN/) RAINES T M.
 XX PA (YUY/) YU Y.
 XX PA (RAME/) RAMEKA J G.
 XX PA (PAGE/) PAGE A.
 XX PA (MATH/) MATHW A V.
 XX PA (LEDF/) LEDFORD B L.
 XX PA (WOES/) WOESSNER J P.
 XX PA (HAAS/) HAAS W D.
 XX PA (GARC/) GARCIA C A.
 XX PA (KRIC/) KRICKER M.
 XX PA (SLAT/) SLATER T.
 XX PA (DAV/) DAVIS K R.
 XX PA (ALLE/) ALLEN K.
 XX PA (HOFF/) HOFFMAN N.
 XX PA (HURB/) HURBAN P.

PI GORLACH J, AN Y, HAMILTON CM, PRICE JL, RAINES TM, YU Y,
 PI RAMEKA JG, PAGE A, MATHW AV, LEDFORD BL, WOESSNER JP, HAAS WD,
 PI GARCIA CA, KRICKER M, SLATER T, DAVIS KR, ALLEN K, HOFFMAN N;
 XX HURBAN P;
 DR WPI; 2003-810930/76.

XX Claim 1; SEQ ID NO 873; 44pp; English.

PT Novel Arabidopsis thaliana nucleic acids useful for generating
 PT genetically modified transgenic organisms, for screening biologically
 PT active agents such as fungicides, insecticides.

CC The invention describes a nucleic acid (I) comprising a sequence capable
 CC of hybridising under stringent conditions to any one of 999 fully defined
 CC Arabidopsis thaliana sequences (I) as given in specification e.g., 360,
 CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a
 CC hybridisation probe to complementary molecules in a cDNA library. (I) is
 CC also useful for generating genetically modified and transgenic organisms,
 CC usually plant cells and plants. A protein encoded by (I) is useful in
 CC screening assays to determine the effect of candidate inhibitors,
 CC activators or modifiers of the gene product. The protein is also useful
 CC for screening biologically active agents e.g., fungicides and
 CC insecticides. A genetically modified cell, comprising an exogenous
 CC nucleic acid, where the nucleic acid comprises transcription regulatory
 CC sequences operably linked to a sequence capable of hybridising under
 CC stringent conditions to (I) is useful in the study of genetic function
 CC and regulation, for alteration of the cellular metabolism and for
 CC screening compounds that may affect the biological function of the gene
 CC or gene product. This sequence represents an Arabidopsis thaliana
 CC polynucleotide of the invention.

SQ Sequence 483 BP; 142 A; 115 C; 76 G; 150 T; 0 U; 0 Other;

Query Match 1.8%; Score 29; DB 9; Length 483;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
 Db 48 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 6

ABZ20190/C

ID ABZ20190 standard; cDNA; 507 BP.

XX AC ABZ20190;

XX DT 23-JAN-2003 (first entry)

XX DE Group III cDNA cancer related clone SEQ ID NO:2616.

XX KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; gene; ss.

XX OS Homo sapiens.

XX PN WO200278516-A2..

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US010421.

XX PR 30-MAR-2001; 2001US-0280255P.

XX PR 28-AUG-2001; 2001US-0315563P.

XX PR 09-JAN-2002; 2002US-0347313P.

XX PA (CORI-) CORIXA CORP.

XX PI Wang T, Wang S, Bangur CS, Gaiger A;

DR WPI; 2003-058387/05.

XX New immunogenic polynucleotides or polypeptides useful for diagnosing, preventing and treating cancer expressing CT or Cp mRNA antigens, and in virology, immunology, microbiology, molecular biology and recombinant DNA techniques.

XX Claim 1; SEQ ID NO 2616; 207pp; English.

XX ABQ17575 to ABQ20506 represent isolated polynucleotide (II) sequences, and ABP54446 to ABP54472 represent protein (II) sequences, from the present invention. (I) and (II) have cytostatic activity and can be used in gene therapy and vaccines. (I), (II), antibodies and compositions from the present invention are useful for diagnosing, preventing and treating cancer, which expresses CT or Cp mRNA antigens. They are useful for stimulating immune response. They can also be useful in virology, immunology, microbiology, molecular biology and recombinant DNA techniques. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 507 BP; 154 A; 79 C; 87 G; 175 T; 0 U; 12 Other;

Query Match 1.8%; Score 29; DB 7; Length 507;

Best Local Similarity 100.0%; Pred. No. 0.1;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602

DB 34 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 6

RESULT 7

AAK64289/c

ID AAK64289 standard; cDNA; 536 BP.

XX AAK64289;

XX DT 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:9349.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX Cytostatic; gene therapy; vaccine; metastasis; ss.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218280P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225575P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 14-AUG-2000; 2000US-0226273P.

PR 18-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227003P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229511P.

PR 06-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 08-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234597P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231144P.
PR 08-SEP-2000; 2000US-0231145P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232197P.
PR 14-SEP-2000; 2000US-0232198P.
PR 14-SEP-2000; 2000US-0232199P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234374P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239393P.
PR 13-OCT-2000; 2000US-0239397P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 05-JAN-2001; 2000US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

P-PSDB; AAU16151.

New nucleic acid molecules encoding 461 human secreted proteins for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives.

Claim 1; SEQ ID NO 317; 980pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention. Note: The sequence data for this patent did not form part of the printed

Query Match 1.8%; Score 29; DB 4; Length 1267;

Best Local Similarity 100.0%; Pred. No. 0.089;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
 DB 1222 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1250

RESULT 9

ABX73479
 ID ABX73479 standard; DNA; 1267 BP.

XX AC ABX73479;

XX DT 18-MAR-2003 (first entry)

XX DE Human novel polynucleotide #307.

XX KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.

XX OS Homo sapiens.

XX FN US2002132753-A1.

XX PD 19-SEP-2002.

XX PP 17-JAN-2001; 2001US-00764864.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 22-AUG-2000; 2000US-0225758P.

XX PR 30-AUG-2000; 2000US-0228688P.

XX PR 01-SEP-2000; 2000US-0229287P.

XX PR 01-SEP-2000; 2000US-0229343P.

XX PR 01-SEP-2000; 2000US-0229344P.

XX PR 01-SEP-2000; 2000US-0229345P.

XX PR 05-SEP-2000; 2000US-0229509P.

XX PR 08-SEP-2000; 2000US-0229513P.

XX PR 21-SEP-2000; 2000US-0231413P.

XX PR 21-SEP-2000; 2000US-0234223P.

XX PR 25-SEP-2000; 2000US-0234274P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241809P.

PR 01-NOV-2000; 2000US-0244617P.

PR 17-NOV-2000; 2000US-0249299P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

XX (RUBE/) RUBEN S M.

XX (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI, 2003-147444/14.

XX P-PSDB; ABUS5219.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,

XX inhibiting or preventing e.g. neural, immune system, muscular,

XX respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or

XX renal disorders.

XX Claim 1; SEQ ID NO 317; 402pp; English.

XX The invention relates to human novel polypeptides and their associated

XX polynucleotides. The polypeptides and polynucleotides are useful in gene

XX therapy for treating, inhibiting or preventing neural disorders, immune

XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis

XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.

XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,

XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,

XX (e.g. congenital heart defects, Eosin's anomaly and hypoplastic left

XX renal syndrome), renal disorders (e.g. acute kidney failure and end-stage

XX leukaemia), hyperproliferative disorders (e.g. Hodgkin's disease and

XX appendicitis), inflammatory diseases (e.g. septic shock, bursitis and

XX related disorders (e.g. thrombosis, atherosclerosis and myocardial

XX infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent

XX human novel polynucleotides of the invention

SQ Sequence 1267 BP; 316 A; 386 C; 317 G; 248 T; 0 U; 0 Other;

Query Match 1.8%; Score 29; DB 7; Length 1267;

Best Local Similarity 100.0%; Pred. No. 0.089;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602

Db 1222 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1250

RESULT 10

ABL86191

ID ABL86191 standard; cDNA; 307 BP.

XX ABL86191;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:9169.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX PA (CORI-) CORIXA CORP.
 XX PI Algate PA, Harlocker SL, Jones R;
 XX XX WPI; 2002-122075/16.
 XX DR
 XX PT Composition for therapy and diagnosis of ovarian cancer comprising
 XX PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 XX PT polypeptide, antibody specific to polypeptide or T cell expressing
 XX PT polypeptide.
 XX PS Claim 1; SEQ ID NO 9169; 489pp; English.
 XX CC The present invention describes a composition (I) comprising: carriers
 XX CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 XX CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
 XX CC from the 10912 nucleotide sequences as given in AB17023 to AB17934,
 XX CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
 XX CC or antigen presenting cells that express (II). (I) has cytostatic
 XX CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
 XX CC detecting ovarian cancer in a patient's biological sample preferably
 XX CC serum or ovarian tissue. The method comprises contacting a biological
 XX CC sample from a patient with (IV), detecting the amount of polynucleotide
 XX CC hybridising to (IV) and comparing the amount to a predetermined cutoff
 XX CC value and thereby detecting ovarian cancer in the patient, where the
 XX CC amount of polynucleotide hybridising to (IV) is detected preferably by
 XX CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
 XX CC useful for stimulating and/or expanding T cells specific for an ovarian
 XX CC tumour protein comprising contacting T cells with (III) or (II). (III) is
 XX CC useful in design and preparation of ribozyme molecules for inhibiting
 XX CC expression of the tumour polypeptides and proteins in tumour cells; and
 XX CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
 XX CC library using well known techniques
 XX SQ Sequence 307 BP; 128 A; 50 C; 62 G; 67 T; 0 U; 0 Other;
 Query Match 1.7%; Score 28; DB 6; Length 307;
 Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;
 Matches 28; Conservative 0;
 QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
 Db 265 TTATGCAAAAAAAAAAAAAAAAAAAAAA 292
 RESULT 11
 AAS60367/C
 ID AAS60367 standard; cDNA; 339 BP.
 AC AAS60367;
 DT 29-JAN-2002 (first entry)
 DE Human cancer agent-resistance marker #241.
 KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
 KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
 KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;
 KW Hodgkin's disease; glioma; sb.
 OS Homo sapiens.
 PN WO200179556-A2.
 PD 25-OCT-2001.
 PF 13-APR-2001; 2001WO-US012132.
 PR 14-APR-2000; 2000US-0197538P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PT

PI Lillie J, Brown JL, Bolt A, Van Ruffel C;
 XX WPI; 2001-602933/68.
 XX PT Novel nucleic acid, used as a marker to determine the effectiveness of
 XX PT using TAXOL to treat cancer cell growth in individuals.
 XX PS Claim 1; Page 199; 527pp; English.
 XX CC The invention relates to 1046 novel nucleic acids which are used as
 XX CC markers for determining the sensitivity of a cancer cell to the
 XX CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when they
 XX CC are shown to express one of the 242 sensitivity markers or the cells are
 XX CC shown not to express one of the 804 resistance markers. The methods can
 XX CC be used to determine the effectiveness of TAXOL in the treatment of
 XX CC cancer cell growth in an individual. The markers can be used as targets
 XX CC in developing anti-cancer agents such as chemotherapeutic compounds. The
 XX CC markers can also be used as targets in developing treatments for cancer,
 XX CC particularly those cancers which display resistance to agents and exhibit
 XX CC expression of the markers. The anticancer agents developed by the novel
 XX CC method can be used to treat cancer. Probes based on the markers can be
 XX CC used to detect transcripts or genomic sequences corresponding to the
 XX CC markers, in the identification of cells or tissues which mis-express the
 XX CC protein. Cancers which may be targeted include carcinoma (e.g. squamous
 XX CC cell carcinoma), sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic
 XX CC leukaemia), lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's
 XX CC disease and tumours (e.g. glioma). The present sequence is one of the
 XX CC 1046 novel cancer cell markers
 XX SQ Sequence 339 BP; 110 A; 73 C; 64 G; 90 T; 0 U; 2 Other;
 Query Match 1.7%; Score 28; DB 4; Length 339;
 Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;
 Matches 28; Conservative 0;
 QY 1573 ATTATGCAAAAAAAAAAAAAAAAAAAAAA 1600
 Db 31 ATTATGCAAAAAAAAAAAAAAAAAAAAAA 4
 RESULT 12
 ABT09151/C
 ID ABT09151 standard; DNA; 419 BP.
 AC ABT09151;
 DT 05-DEC-2002 (first entry)
 DE Phase-1 Rat CT gene SEQ ID No 239.
 DE Rat; toxicity study; rat toxic response gene; toxicological response;
 KW drug development; phase-1 rat CT gene; db.
 OS Rattus sp.
 PN WO200266682-A2.
 PD 29-AUG-2002.
 PF 29-JAN-2002; 2002WO-US002935.
 PR 29-JAN-2001; 2001US-0264933P.
 PR 26-JUL-2001; 2001US-0308161P.
 PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
 PI Farris G, Hicken SH, Farr SB;
 DR WPI; 2002-674961/72.
 XX PT Evaluating the toxicity of an agent; useful in drug development or in
 XX PT determining toxicological responses to a new drug by determining the
 XX PT expression of rat toxicologically relevant genes in the test animal in

PT response to the test agent.

PS Disclosure; Page 165; 388pp; English.

XX The invention relates to a method used for evaluating the toxicity of an agent comprising determining the expression of a rat toxic response gene(s) in the test animal in response to the agent. The method is useful in drug development, particularly for conducting toxicity studies and analysis before a new drug or compound is approved for human consumption or use. The method is also useful in determining toxicological responses to a new drug. This polynucleotide sequence represents a phase-1 rat CT gene of the invention

XX Sequence 419 BP; 153 A; 89 C; 85 G; 92 T; 0 U; 0 Other;

Query Match 1.7%; Score 28; DB 6; Length 419;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1601
|||||
Db 28 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 13
ABA57560/c
ID ABA57560 standard; DNA; 450 BP.
XX
AC ABA57560;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #5865.
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver.
XX
XX Claim 1; SEQ ID NO 5865; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human fetal liver. The present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 450 BP; 123 A; 82 C; 95 G; 150 T; 0 U; 0 Other;

Query Match 1.7%; Score 28; DB 6; Length 419;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
|||||
Db 85 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 58

RESULT 14
AAI37117/c
ID AAI37117 standard; DNA; 450 BP.
XX
AC AAI37117;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #5803 used to measure gene expression in human placenta sample.
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta.
XX
XX Claim 25; SEQ ID NO 5803; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP). The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders
XX
XX Sequence 450 BP; 123 A; 82 C; 95 G; 150 T; 0 U; 0 Other;

Query Match 1.7%; Score 28; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
|||||
Db 85 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 58

RESULT 15
ABA27022/c
ID ABA27022 standard; DNA; 450 BP.
XX
XX ABA27022;
XX

DT 23-JAN-2002 (first entry)
XX Probe #5488 for gene expression analysis in human heart cell sample.
DE Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
XX Claim 1; SEQ ID NO 5488; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 450 BP; 123 A; 82 C; 95 G; 150 T; 0 U; 0 Other;
Query Match 1.7%; Score 28; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. NO. 0.26;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 85 TTATGCAAAAAAAAAAAAAAAAAAAAAA 58

Search completed: August 2, 2004, 11:24:58
Job time : 672 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 08:46:58 ; Search time 176 Seconds

(without alignments)
5051.314 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602

Sequence: 1 cacacgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	1.7	233	4	US-09-621-976-10675
2	27	1.7	1319	4	US-09-443-067-17
3	26	1.6	32	4	US-09-619-103-14
4	26	1.6	38	4	US-09-619-103-2
5	26	1.6	249	4	US-09-621-976-19144
6	26	1.6	787	4	US-09-621-976-1878
7	26	1.6	971	3	US-09-248-335-65
8	26	1.6	1181	4	US-09-149-476-310
9	26	1.6	1195	1	US-08-373-858-1
10	26	1.6	1195	1	US-08-342-786B-1
11	26	1.6	1212	4	US-09-149-476-186
12	26	1.6	1359	3	US-09-387-574-11
13	26	1.6	1359	4	US-09-668-096-11
14	26	1.6	1534	1	US-08-300-903A-6
15	26	1.6	1534	4	US-08-988-197-6
16	26	1.6	1946	4	US-09-882-835-1
17	26	1.6	2971	4	US-09-482-273-37
18	26	1.6	3275	4	US-09-370-838-151
19	26	1.6	44848	4	US-09-435-739-42
20	26	1.6	90050	3	US-09-245-041-5
21	25	1.6	57	1	US-08-120-827-93
22	25	1.6	57	1	US-08-478-675-93
23	25	1.6	57	3	US-09-199-737-59
24	25	1.6	57	4	US-09-058-333A-59
25	25	1.6	104	4	US-09-621-976-11198
26	25	1.6	114	4	US-09-621-976-9602
27	25	1.6	173	4	US-09-621-976-18333

28 25 1.6 191 4 US-09-621-976-18332 Sequence 18332, A

29 25 1.6 261 4 US-09-621-976-17608 Sequence 17608, A

30 25 1.6 269 4 US-09-621-976-17610 Sequence 17610, A

31 25 1.6 348 4 US-09-621-976-18665 Sequence 18665, A

32 25 1.6 362 4 US-09-621-976-10457 Sequence 10457, A

33 25 1.6 455 4 US-09-227-357-70 Sequence 70, Appl

34 25 1.6 462 3 US-09-385-982-504 Sequence 504, App

35 25 1.6 610 3 US-09-385-982-229 Sequence 229, App

36 25 1.6 625 3 US-09-385-982-242 Sequence 242, App

37 25 1.6 629 3 US-09-385-982-389 Sequence 389, App

38 25 1.6 651 1 US-08-171-385-4 Sequence 4, Appl

39 25 1.6 655 4 US-09-288-143-20 Sequence 20, Appl

40 25 1.6 680 4 US-08-865-297-5 Sequence 5, Appl

41 25 1.6 770 3 US-08-938-675A-1 Sequence 1, Appl

42 25 1.6 770 4 US-09-531-727-1 Sequence 1, Appl

43 25 1.6 857 1 US-08-308-883-1 Sequence 1, Appl

44 25 1.6 857 1 US-08-730-163-1 Sequence 1, Appl

45 25 1.6 857 1 US-08-730-163-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-621-976-10675
; Sequence 10675, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTe and Encoded Human Proteins.

; FILE REFERENCE: GENSET 054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT PILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 10675

; LENGTH: 233

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-10675

Query Match 1.7%; Score 27; DB 4; Length 233;
Best Local Similarity 100.0%; Pred. No. 0.008;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1576 TATGCAAAAAAAAAAAAAAAAAAAAAA 1602

Db 178 TATGCAAAAAAAAAAAAAAAAAAAAAA 204

RESULT 2

US-09-443-067-17

; Sequence 17, Application US/09443067

; Patent No. 6627794

; GENERAL INFORMATION:

; APPLICANT: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH

; APPLICANT: ORGANISATION

; TITLE OF INVENTION: Polyphenol oxidase genes from banana, lettuce, tobacco and

; FILE REFERENCE: pineapple

; CURRENT APPLICATION NUMBER: US/09/443,067

; CURRENT PILING DATE: 1999-11-18

; EARLIER APPLICATION NUMBER: US 08/976, 222

; EARLIER PILING DATE: 1997-11-21

; EARLIER APPLICATION NUMBER: PCT/AU98/00362

; EARLIER PILING DATE: 1998-05-19

; EARLIER APPLICATION NUMBER: AU PP3698

; EARLIER PILING DATE: 1995-05-23

; EARLIER APPLICATION NUMBER: AU PP6849

; EARLIER PILING DATE: 1997-05-19

; EARLIER APPLICATION NUMBER: AU PP5600

;; EARLIER FILING DATE: 1995-09-26
;; NUMBER OF SEQ ID NOS: 49
;; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: pineapple
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1053)
US-09-443-067-17

Query Match 1.7%; Score 27; DB 4; Length 1319;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1576 TATCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 1288 TATCAAAAAAAAAAAAAAAAAAAAAA 1314

RESULT 3
US-09-619-103-14
; Sequence 14, Application US/09619103
; Patent No. 6429300
; GENERAL INFORMATION:
; APPLICANT: Kurz, Markus
; APPLICANT: Lohse, Peter
; APPLICANT: Wagner, Richard
; TITLE OF INVENTION: Peptide Acceptor Ligation Methods
; FILE REFERENCE: 50036/031002
; CURRENT APPLICATION NUMBER: US/09/619,103
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/145,834
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 32
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: designed sequence to act as a linker
US-09-619-103-14

Query Match 1.6%; Score 26; DB 4; Length 32;
Best Local Similarity 96.2%; Pred. No. 0.024;
Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCACAAAAAAAAAAAAAAAAAAAA 1602
DB 7 AUGCAAAAAAAAAAAAAAAAAAAAAA 32

RESULT 4
US-09-619-103-2
; Sequence 2, Application US/09619103
; Patent No. 6429300
; GENERAL INFORMATION:
; APPLICANT: Kurz, Markus
; APPLICANT: Lohse, Peter
; APPLICANT: Wagner, Richard
; TITLE OF INVENTION: Peptide Acceptor Ligation Methods
; FILE REFERENCE: 50036/031002
; CURRENT APPLICATION NUMBER: US/09/619,103
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/145,834
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 38
; TYPE: DNA

;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: designed sequence to act as a linker
US-09-619-103-2

Query Match 1.6%; Score 26; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCACAAAAAAAAAAAAAAAAAAAA 1602
DB 6 ATGCACAAAAAAAAAAAAAAAAAAAA 31

RESULT 5
US-09-621-976-19144
; Sequence 19144, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 19144
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-19144

Query Match 1.6%; Score 26; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCACAAAAAAAAAAAAAAAAAAAA 1602
DB 188 ATGCACAAAAAAAAAAAAAAAAAAAA 213

RESULT 6
US-09-621-976-1878
; Sequence 1878, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1878
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 152..673
; NAME/KEY: sig_peptide
; LOCATION: 152..208
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.09999990463257
; OTHER INFORMATION: seq LLIGQRCSLKVG/QE
US-09-621-976-1878

Query Match 1.6%; Score 26; DB 4; Length 787;
Best Local Similarity 100.0%; Pred. No. 0.022;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 699 ATGCAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 7

US-09-248-335-65
; Sequence 65, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 65
; LENGTH: 971
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-65

Query Match 1.6%; Score 26; DB 3; Length 971;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 940 ATGCAAAAAAAAAAAAAAAAAAAAAA 965

RESULT 8

US-09-149-476-310
; Sequence 310, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 1.6%; Score 26; DB 4; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.022; 0; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;
QY 1577 ATGCACAAAAA 1602
DB 1123 ATGCACAAAAA 1148

RESULT 9

US-08-373-858-1
Sequence 1, Application US/08373858
Patent No. 5633155
GENERAL INFORMATION:
APPLICANT: Kim, Man-Keun
APPLICANT: Lee, Kwan-Ho
APPLICANT: Na, Byeong-Kook
APPLICANT: Jeong, Han-Seung
APPLICANT: Choo, Kyu-Whan
APPLICANT: Moon, Young-Ho
APPLICANT: Jeon, Hong-Seob
TITLE OF INVENTION: Expression Vector for Phytoacca
TITLE OF INVENTION: Antiviral Protein and Process for Preparing Transgenic
TITLE OF INVENTION: Plant Transformed Thereof.
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: US
ZIP: 10022-7513
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/373,858
FILING DATE: 18-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Peter
REGISTRATION NUMBER: 25,351
REFERENCE/DOCKET NUMBER: 0136/17986-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7770
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1195 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: *Phytolacca americana*
TISSUE TYPE: Leaf
US-08-373-858-1

Query Match 1.6%; Score 26; DB 1; Length 1195;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels

Qy 1577 ATGCAAAAAAAAAAAAAAAAAA 1602
|||
Db 1169 ATGCAAAAAAAAAAAAAAAAAA 1194

RESULT 10
US-08-342-786B-1
; Sequence 1, Application US/08342786B
; Patent No. 5648234
; GENERAL INFORMATION:
; APPLICANT: Moon, Young-Ho
; APPLICANT: Jeon, Hong-Seob
; APPLICANT: Choi, Kyu-Whan
; APPLICANT: Lee, Kwan-Ho
; APPLICANT: Kim, Man-Keun
; TITLE OF INVENTION: A NO. 5648234
; TITLE OF INVENTION: Antiviral pro
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.

US-08-342-786B-1

Query Match 1.6%; Score 26; DB 1; Length 1195;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels

Qy	1577	ATGCAAAAAAAAAAAAAAAAAA	1602
Db	1169	ATGCAAAAAAAAAAAAAAAAAA	1194

RESULT 11

US-09-149-476-186

Sequence 186, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

CURRENT APPLICATION NUMBER: US/09/149,476

CURRENT FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,592

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,587

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,492

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,582

EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 1.68; Score 26; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1577 ATGCACAAAAA 1602

OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word for Apple, Version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,197
FILING DATE:
CLASSIFICATION: 121097
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/300,903
FILING DATE: 06-SEPTEMBER-1994
APPLICATION NUMBER: USSN 08/236,919
FILING DATE: 06-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2822-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
TELEFAX: 206-233-0644
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1534 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..753
US-08-988-197-6

Query Match 1.6%; Score 26; DB 4; Length 1534;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1577 ATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 1476 ATGCAAAAAAAAAAAAAAAAAAAAAA 1501

Search completed: August 2, 2004, 14:51:07
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 13:13:14 ; Search time 767 Seconds
(without alignments)

10240.978 Million cell updates/sec

Title: US-10-031-331b-39

Perfect score: 1602
Sequence: 1 cacacgttgattaccat.....aaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 3222919 seqs, 2451570024 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30	1.9	504	13	US-10-424-599-121946
2	30	1.9	523	17	US-10-021-323-11875
3	29	1.8	416	13	US-10-424-599-71220
C 4	29	1.8	483	10	US-09-770-961-873
5	29	1.8	496	17	US-10-021-323-4109
6	29	1.8	1267	9	US-09-764-864-317
C 7	29	1.8	43853	17	US-10-322-281-788
8	28	1.7	307	9	US-09-867-701-9169
C 9	28	1.7	339	9	US-09-834-975-368
C 10	28	1.7	450	9	US-09-864-761-5488
11	28	1.7	559	13	US-10-342-587-1027
12	28	1.7	559	13	US-10-172-118-1077
13	28	1.7	559	13	US-10-170-385-468
14	28	1.7	809	17	US-10-437-963-87473

C 15 28 1.7 1021 17 US-10-437-963-47572
16 28 1.7 1051 13 US-10-424-599-140317
17 28 1.7 1069 9 US-10-437-963-30744
18 28 1.7 2000 17 US-09-938-842A-4789
19 28 1.7 2000 11 US-09-938-842A-4789
20 28 1.7 3183 15 US-10-442-017-8
21 28 1.7 45121 12 US-09-997-722-10
22 28 1.7 45121 12 US-09-997-722-10
C 23 27 1.7 43 13 US-10-175-539A-45
24 27 1.7 159 13 US-10-085-783A-14143
25 27 1.7 159 16 US-10-242-535A-14143
26 27 1.7 424 15 US-10-198-846-13129
C 27 27 1.7 444 9 US-09-962-436-88
28 27 1.7 444 9 US-09-880-107-2880
C 29 27 1.7 444 9 US-09-967-768A-167
30 27 1.7 474 13 US-10-027-632-49689
C 31 27 1.7 474 16 US-10-027-632-49689
C 32 27 1.7 491 12 US-09-969-034-984
33 27 1.7 492 17 US-10-437-963-3082
C 34 27 1.7 505 13 US-10-424-599-45864
35 27 1.7 513 10 US-09-918-995-21477
36 27 1.7 542 13 US-10-027-632-5404
37 27 1.7 542 16 US-10-027-632-5404
C 38 27 1.7 681 15 US-10-086-543-976
39 27 1.7 854 15 US-10-198-846-6097
40 27 1.7 979 16 US-10-159-563-298
41 27 1.7 1164 17 US-10-437-963-28218
42 27 1.7 1170 16 US-10-264-049-927
43 27 1.7 1201 13 US-10-342-887-955
44 27 1.7 1201 13 US-10-172-118-955
45 27 1.7 1335 9 US-09-925-300-534

ALIGNMENTS

RESULT 1
US-10-424-599-121946/c
; Sequence 121946, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 121946
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81123C.1
US-10-424-599-121946

Query Match 1.9%; Score 30; DB 13; Length 504;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1573 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 44 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 15

RESULT 2

US-10-021-323-11875
; Sequence 11875, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:

APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd B.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021.323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 11875
LENGTH: 523
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3829-014-Q6-K6-A4
US-10-021-323-11875

Query Match 1.9%; Score 30; DB 17; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1573 ATTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 307 ATTATGCAAAAAAAAAAAAAAAAAAAAAA 336

RESULT 3
US-10-424-599-71220
Sequence 71220, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 71220
LENGTH: 416
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT MRT3847_35323C.1
US-10-424-599-71220

Query Match 1.8%; Score 29; DB 13; Length 416;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 144 TTATGCAAAAAAAAAAAAAAAAAAAAAA 172

RESULT 4
US-09-770-961-873/c
Sequence 873, Application US/09770961
Publication No. US20030115639A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Giang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy
APPLICANT: Matthew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Naja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurlban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2026 (PARA-015PRV)
CURRENT APPLICATION NUMBER: US/09/770,961
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/179,466
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 873
LENGTH: 483
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-961-873

Query Match 1.8%; Score 29; DB 10; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 48 TTATGCAAAAAAAAAAAAAAAAAAAAAA 20

RESULT 5
US-10-021-323-4109
Sequence 4109, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd B.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 4109
LENGTH: 496
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(496)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3825-031-Q6-K6-B6
US-10-021-323-4109

Query Match 1.8%; Score 29; DB 17; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 194 TTATGCAAAAAAAAAAAAAAAAAAAAAA 222

RESULT 6

US-09-764-864-317
; Sequence 317, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 317
; LENGTH: 1267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-317

Query Match 1.8%; Score 29; DB 9; Length 1267;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 1222 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1250

RESULT 7

US-10-322-281-788/c
; Sequence 788, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 788
; LENGTH: 43853
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-788

Query Match 1.8%; Score 29; DB 17; Length 43853;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1574 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 40399 TTTATGCAAAAAAAAAAAAAAAAAAAAAA 40371

RESULT 8

US-09-867-701-9169
; Sequence 9169, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9169
; LENGTH: 307

; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9169

Query Match 1.7%; Score 28; DB 9; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
Db 265 TTATGCAAAAAAAAAAAAAAAAAAAAAA 292

RESULT 9

US-09-834-975-368/c
; Sequence 368, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 368
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(339)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-368

Query Match 1.7%; Score 28; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1573 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 1600
Db 31 ATTTATGCAAAAAAAAAAAAAAAAAAAAAA 4

RESULT 10

US-09-864-761-5488/c
; Sequence 5488, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US/09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 5488
LENGTH: 450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC005479.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
US-09-864-761-5488

Query Match 1.7%; Score 28; DB 9; Length 450;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 85 TTATGCAAAAAAAAAAAAAAAAAAAAAA 58

RESULT 11

US-10-342-887-1077
Sequence 1077, Application US/10342887.
Publication No. US20040058340A1
GENERAL INFORMATION:
APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Linsley, Peter S.
APPLICANT: Mao, Mao
APPLICANT: Roberts, Christopher J.
APPLICANT: Van 't Veer, Laura Johanna
APPLICANT: Van de Vijver, Marc J.
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-188-999
CURRENT APPLICATION NUMBER: US/10/342,887
PRIOR FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/298,918
PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/380,710
PRIOR FILING DATE: 2002-05-14
PRIOR APPLICATION NUMBER: 10/172,118
PRIOR FILING DATE: 2002-06-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1077
LENGTH: 559
TYPE: DNA
ORGANISM: Homo sapiens
US-10-342-887-1077

Query Match 1.7%; Score 28; DB 13; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 513 TTATGCAAAAAAAAAAAAAAAAAAAAAA 540

RESULT 12

US-10-172-118-1077
Sequence 1077, Application US/10172118
Publication No. US20030224374A1
GENERAL INFORMATION:
APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Linsley, Peter
APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT APPLICATION NUMBER: US/10/172,118
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1077
LENGTH: 559
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NM_004708
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1077

Query Match 1.7%; Score 28; DB 13; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 513 TTATGCAAAAAAAAAAAAAAAAAAAAAA 540

RESULT 13

US-10-170-385-468
Sequence 468, Application US/10170385
Publication No. US20030203372A1
GENERAL INFORMATION:
APPLICANT: Ward, Neil Raymond
APPLICANT: Mundy, Christopher Robert
APPLICANT: Kan, On
APPLICANT: Harris, Robert Alan
APPLICANT: White, Jonathan
APPLICANT: Binley, Katie Mary
APPLICANT: Rayner, William Nigel
APPLICANT: Naylor, Stuart
APPLICANT: Kingsman, Susan Mary
APPLICANT: Krige, David


```
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 468
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-170-385-468
```

```
Query Match 1.7%; Score 28; DB 13; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 513 TTATGCAAAAAAAAAAAAAAAAAAAAAA 540
```

RESULT 14

```
US-10-437-963-87473
; Sequence 87473; Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87473
; LENGTH: 809
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86414C.1
; US-10-437-963-87473
```

```
Query Match 1.7%; Score 28; DB 17; Length 809;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 780 TTATGCAAAAAAAAAAAAAAAAAAAAAA 807
```

RESULT 15

```
US-10-437-963-47572/c
; Sequence 47572; Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
```

```
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 47572
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50327C.1
; US-10-437-963-47572
```

```
Query Match 1.7%; Score 28; DB 17; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1575 TTATGCAAAAAAAAAAAAAAAAAAAAAA 1602
DB 36 TTATGCAAAAAAAAAAAAAAAAAAAAAA 9
```

Search completed: August 2, 2004, 16:27:05
Job time : 774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2004, 11:13:45 ; Search time 4443 Seconds
(without alignments)
10767.322 Million cell updates/sec

Title: US-10-031-331B-39
Perfect score: 1602
Sequence: 1 cacacgttgatttaaccat.....aaaaaaaaaaaaaaaaaaaa 1602

Scoring table: OLIGO NUC
Gapop_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST.*

1: em_estba.*
2: em_estbam.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_pbg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	368	23.0	713..10..	BE231445	S
2	35	2.2	552	10	BF051218
3	35	2.2	568	10	BF051675
4	35	2.2	583	14	CB256828

5	35	2.2	593	12	BM113004
6	35	2.2	602	10	AW623904
7	35	2.2	602	13	BQ114297
8	35	2.2	604	10	BE913890
9	35	2.2	605	10	AW621907
10	35	2.2	605	10	AW623370
11	35	2.2	629	10	AW649895
12	35	2.2	654	12	BG642645
13	35	2.2	722	12	BG128609
14	35	2.2	722	12	BI931842
15	32	2.0	595	14	CA847413
16	32	2.0	603	12	BUS55248
17	32	2.0	605	12	BUS562413
18	32	2.0	703	14	CB342962
19	31	1.9	458	12	BI841667
20	31	1.9	518	12	BI396184
21	31	1.9	560	12	BUS62806
22	31	1.9	560	13	BQ718025
23	31	1.9	608	14	CD482486
24	30	1.9	122	9	A2771389
25	30	1.9	124	9	A1611367
26	30	1.9	126	9	A1252727
27	30	1.9	127	9	A1250129
28	30	1.9	127	9	A1802850
29	30	1.9	128	9	A1371930
30	30	1.9	129	9	A1312227
31	30	1.9	129	9	A1312386
32	30	1.9	144	9	A1345222
33	30	1.9	146	9	A1345215
34	30	1.9	150	9	A1305426
35	30	1.9	150	9	A1305559
36	30	1.9	150	9	A1263547
37	30	1.9	156	14	CD801054
38	30	1.9	164	9	A1266670
39	30	1.9	232	13	BUS35860
40	30	1.9	237	14	CD642043
41	30	1.9	247	13	BQ800441
42	30	1.9	315	13	BU780369
43	30	1.9	336	10	AW682563
44	30	1.9	340	10	BF014794
45	30	1.9	388	9	A1666802

ALIGNMENTS

BE231445 713 bp mRNA linear EST 16-OCT-2000
Ss0815 Suaeda salsa ZAP cDNA library Suaeda maritima subsp. salsa
CDNA similar to c.elegans CDNA y492b11.3, mRNA sequence.

BE231445.1 GI:8996113
EST.

Suaeda maritima subsp. salsa
Suaeda maritima subsp. salsa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Suaeda.

1 (bases 1 to 713)

Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.

Expressed sequence tags from a halophyte Suaeda salsa CDNA library

Unpublished (2000)

Contact: Hui Zhang

Key Laboratory of Plant Stress Research

The Biology Department of Shandong Normal University

No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC

Tel: (86)531-2960864

Fax: (86)531-2966954

Email: zhangh@sndu.edu.cn.

Location/Qualifiers

1. 713

/organism="Suaeda maritima subsp. salsa"

FEATURES
source


```

Query Match      2.2%; Score 35; DB 10; Length 568;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 278
      |||||
DB 487 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 521

RESULT 4
LOCUS CB256828
DEFINITION 58-E012740-027-002-D16-T7R MP12-ADIS-027 Arabidopsis thaliana cDNA
ACCESSION CB256828
VERSION CB256828.1 GI:32881601
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 583)
AUTHORS Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
Mitchell-Olds,T. and Weishaar,B.
TITLE Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE 22683290
PUBMED 12799357
COMMENT Contact: Weishaar B
ADIS DNA core facility at MP12
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishaar@mpiz-koeln.mpg.de
Insert Length: 583 Std Error: 0.00
Plate: 2 row: D column: 16
Seq primer: T7R; CTAATAGACTCTACTATAGGGA.

FEATURES
source
1..583
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Eifel-2 (Si-2)"
/db_xref="GABI:593271"
/db_xref="taxon:3702"
/clone="MP12p772D162Q"
/tissue_type="whole plant"
/dev_stages="adult plant, mixed stresses"
/lab_host="S. coli TOP10"
/clone_lib="MP12-ADIS-027"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Arabidopsis thaliana, accession Eifel-2; ten
week old total plants grown under long-day conditions in
soil, whole adult plants were treated for 24 hours with
different stresses, (1) at 4M-0 C in the dark, (2), at 37
Grad C in the dark, (3) lying in the lab after removing
from soil, (4) in the greenhouse after wounding leaves
with a forceps, (5) in the lab watering with a 150 mM
NaCl solution, (6) at 26 M-0C in the light/UV; equal
quantities of stressed plant material were pooled; library
was made at the Max-Planck-Institute for Plant Breeding
Research, Cologne, Germany; cloning sites SalI-NotI,
primer sites and orientation:
T7-SalI-CCAGCGTCGCG-5prime-cDNA-polyA-CC-NotI-Sp6; Note:
Sequencing granted in the context of the GABI Arabidopsis
Verbund I: Genetic Diversity, 'Establishment of
high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection' PI: Bernd
Weishaar Sequence submission-managed-by-RZPD/GABI-Primary
database: http://gabi.rzpd.de. This clone is available
from RZPD; contact RZPD (clone@rzpd.de) for further
information."

```

```

ORIGIN
Query Match      2.2%; Score 35; DB 14; Length 583;
Best Local Similarity 100.0%; Pred. No. 2.2e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AAGCTTCTGATCTTGACAAAGAGAACGCTCTGAG 78
      |||||
DB 313 AAGCTTCTGATCTTGACAAAGAGAACGCTCTGAG 347

RESULT 5
LOCUS BM113004
DEFINITION EST560540 potato roots Solanum tuberosum cDNA clone CPRO17C13 5'
end, mRNA sequence.
ACCESSION BM113004
VERSION BM113004.1 GI:17076052
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 593)
AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,
Utterback,T., Chiemingo,A., Bougri,O., Buell,C.R., Romning,C.,
Tankaley,S. and Baker,B.
TITLE Generation of ESTs from potato roots
JOURNAL Unpublished (2001)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3

FEATURES
source
1..593
Location/Qualifiers
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CPRO17C13"
/tissue_type="roots"
/dev_stages="in vitro grown stem cuttings"
/lab_host="SOLR"
/clone_lib="potato roots"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

ORIGIN
Query Match      2.2%; Score 35; DB 12; Length 593;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 244 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 278
      |||||
DB 475 AATGTCAGTTATGTCGCTGATGCTGACTTCTCC 509

RESULT 6
LOCUS AW623904
DEFINITION EST121849 tomato flower buds 3-8 mm, Cornell University
Lycopersicon esculentum cDNA clone CTOB13B5 5', mRNA sequence.
ACCESSION AW623904
VERSION AW623904.1 GI:7336931

```

ORIGIN

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278
 Db 500 AATGTCACGTTATGTCGTCGATGACTTCCTCC 534

RESULT 9
 AW621907
 LOCUS
 DEFINITION EST132705 tomato root during/after fruit set, Cornell University
 Lycopersicon esculentum cDNA clone cLEX13N23 5', mRNA sequence.
 AW621907
 VERSION
 SOURCE
 ORGANISM
 Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 605)
 van der Hoeven, R.S., Garvin, D., Matern, A.L., Holt, I.E., Kochian, L.,
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and
 Tanksley, S.D.
 Generation of ESTs from tomato root, during and after fruit set
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..605
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLEX13N23"
 /tissue_type="root"
 /dev_stages="plants during and after fruit-set"
 /clone_lib="tomato root during/after fruit set, Cornell
 University"
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin
 (USDA-ARS, Ithaca, NY 14850)."

ORIGIN
 Query Match 2.2%; Score 35; DB 10; Length 605;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 10
 AW623370
 LOCUS
 DEFINITION EST132315 tomato flower buds 3-8 mm, Cornell University
 Lycopersicon esculentum cDNA clone cTOB10A19 5', mRNA sequence.
 AW623370
 VERSION
 SOURCE
 ORGANISM
 Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 605)
 van der Hoeven, R.S., Bezzereades, J.L., Matern, A.L., Holt, I.E.,

Query Match 2.2%; Score 35; DB 10; Length 605;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11
 AW649895
 LOCUS
 DEFINITION EST1328349 tomato germinating seedlings, TAMU Lycopersicon
 esculentum cDNA clone cLE111C11 5', mRNA sequence.
 AW649895
 VERSION
 SOURCE
 ORGANISM
 Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 629)
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and
 Tanksley, S.D.
 Generation of ESTs from germinating tomato seed
 Unpublished (2000)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..629
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLE111C11"
 /tissue_type="whole seedlings"
 /dev_stage="7 days post imbibition"
 /clone_lib="tomato germinating seedlings, TAMU"

Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and
 Tanksley, S.D.
 Generation of ESTs from tomato flower tissue, 3-8 mm buds
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..605
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cTOB10A19"
 /tissue_type="flower"
 /dev_stages="3-8mm buds"
 /clone_lib="tomato flower buds 3-8 mm, Cornell University"
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Tanksley; Flower buds and flowers were
 taken from greenhouse plants (4-8 wks old, TA496). They
 were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

ORIGIN
 Query Match 2.2%; Score 35; DB 10; Length 605;
 Best Local Similarity 100.0%; Pred. No. 2.1e+03;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 244 AATGTCACGTTATGTCGTCGATGACTTCCTCC 278
 Db 498 AATGTCACGTTATGTCGTCGATGACTTCCTCC 532

RESULT 11
 AW649895
 LOCUS
 DEFINITION EST1328349 tomato germinating seedlings, TAMU Lycopersicon
 esculentum cDNA clone cLE111C11 5', mRNA sequence.
 AW649895
 VERSION
 SOURCE
 ORGANISM
 Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 629)
 Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,
 Nierman, W., Fraser, C.M., Giovannoni, J.J., Martin, G.B. and
 Tanksley, S.D.
 Generation of ESTs from germinating tomato seed
 Unpublished (2000)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..629
 Location/Qualifiers
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLE111C11"
 /tissue_type="whole seedlings"
 /dev_stage="7 days post imbibition"
 /clone_lib="tomato germinating seedlings, TAMU"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; 7 days post imbibition on water-agar. Mixed stage whole germinating seedlings from seed coat emergence up to two centimeters in seeds not showing obvious signs of germination were discarded."

ORIGIN

Query Match 2.2%; Score 35; DB 10; Length 629;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 244 AATGTCAGTTTATGTGCTGATGACTTCTCC 278
|||||
Db 482 AATGTCAGTTTATGTGCTGATGACTTCTCC 516
|||||

RESULT 12

BG642645 654 bp mRNA linear EST 24-APR-2001
LOCUS EST510839 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION CTOF25C16 5' sequence, mRNA sequence.

ACCESSION BG642645

VERSION 1 GI:13777498

KEYWORDS EST

SOURCE Lycopersicon esculentum (tomato)

ORGANISM

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 654)

REFERENCE van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Utterback, T.,

Generation of ESTs from tomato shoot/meristem tissue

Unpublished (2001)

JOURNAL

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES

source

1. 654

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CTOF25C16"

/tissue_type="shoot/meristem"

/dev_stage="developing shoots from 4-6wks old plants"

/lab_host="SOLR"

/clone_lib="tomato shoot/meristem"

/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; Small expanding leaves from the growing tip were

taken from greenhouse plants (4-6wks old TA496). Tissue

was immediately frozen in liquid nitrogen."

ORIGIN

Query Match 2.2%; Score 35; DB 12; Length 654;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 244 AATGTCAGTTTATGTGCTGATGACTTCTCC 278
|||||

Db 318 AATGTCAGTTTATGTGCTGATGACTTCTCC 352
|||||

RESULT 13

BG128609 722 bp mRNA linear EST 31-JAN-2001

LOCUS EST474255 tomato shoot/meristem Lycopersicon esculentum cDNA clone

DEFINITION CTOF21D11 5' sequence, mRNA sequence.

ACCESSION BG128609

VERSION BG128609.1 GI:12628797

KEYWORDS EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

Location/Qualifiers

1. 722

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CTOF21D11"

/tissue_type="shoot/meristem"

/dev_stage="developing shoots from 4-6wks old plants"

/lab_host="SOLR"

/clone_lib="tomato shoot/meristem"

/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; Small expanding leaves from the growing tip were

taken from greenhouse plants (4-6wks old TA496). Tissue

was immediately frozen in liquid nitrogen."

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Oy

Db

244

481

AATGTCAGTTTATGTGCTGATGACTTCTCC 278

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

AATGTCAGTTTATGTGCTGATGACTTCTCC 515

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.